Commun. Math. Phys. 359, 347–373 (2018)
Digital Object Identifier (DOI) https://doi.org/10.1007/s00220-017-2999-2

Communications in
**Mathematical
Physics**

Lyapunov Exponents and Correlation Decay for Random Perturbations of Some Prototypical 2D Maps

Alex Blumenthal[1], **Jinxin Xue**[2], **Lai-Sang Young**[3]

[1] Department of Mathematics, University of Maryland, College Park, MD, USA.
E-mail: alexb123@math.umd.edu
[2] Department of Mathematics, University of Chicago, Chicago, IL, USA.
E-mail: jxue@math.uchicago.edu
[3] Courant Institute of Mathematical Sciences, New York University, New York, NY, USA.
E-mail: lsy@cims.nyu.edu

Received: 4 March 2017 / Accepted: 1 August 2017
Published online: 12 October 2017 –

Abstract: To illustrate the more tractable properties of random dynamical systems, we consider a class of 2D maps with strong expansion on large—but non-invariant—subsets of their phase spaces. In the deterministic case, such maps are not precluded from having sinks, as derivative growth on disjoint time intervals can be cancelled when stable and unstable directions are reversed. Our main result is that when randomly perturbed, these maps possess positive Lyapunov exponents commensurate with the amount of expansion in the system. We show also that initial conditions converge exponentially fast to the stationary state, equivalently time correlations decay exponentially fast. These properties depend only on finite-time dynamics, and do not involve parameter selections, which are necessary for deterministic maps with nonuniform derivative growth.

Two signatures of chaotic behavior in dynamical systems are the positivity of Lyapunov exponents and fast decay of time correlations. For deterministic maps that are not uniformly expanding or hyperbolic, these properties can be difficult to prove even when the underlying geometry suggests a strong likelihood of chaotic behavior. Our main message is that the situation for *random maps* is different, and one of the aims of this paper is to propose a systematic way to establish the positivity of Lyapunov exponents for such maps. For a prototypical class of 2D maps with certain requisite geometry, we recover, under small perturbations, Lyapunov exponents that correctly reflect that geometry. We prove also that the time correlations of such maps decay exponentially fast.

We begin with a discussion of the underlying issues before proceeding to a more detailed discussion of our results.

Alex Blumenthal: This research was supported by NSF Grant DMS-1604805. Jinxin Xue: This research was supported by NSF Grant DMS-1500897. Lai-Sang Young: This research was supported in part by NSF Grant DMS-1363161.

The situation vis a vis deterministic maps

Proving rigorously the chaotic behavior of a deterministic map F entails the following challenges: Even if $\|dF\| \gg 1$ on a large part of the phase space, if F does not possess a continuous family of invariant cones, i.e., if its expanding and contracting directions are not well separated, then the sequence $\|dF_x^n(v)\|$, $n = 1, 2, \ldots$, for a typical point x and vector v can oscillate with n, and whether its cumulative growth rate is positive or negative can be a delicate cancellation problem. These cancellations are real; see e.g. [5,6,10]. The challenge of estimating the asymptotic growth rates of $\|dF_x^n(v)\|$ is exemplified by the Chirikov standard map [8], for which it is not known whether a positive Lyapunov exponent exists on a positive Lebesgue measure set of x for arbitrarily large expansion constants; see e.g. [11].

If F is not volume-preserving, and one is interested in physically relevant observations, then additionally one will have to prove the existence of an SRB measure. (These issues are discussed in the review article [29], Sect. 5.) As for the rate of mixing, current techniques require properties considerably more stringent than just positive Lyapunov exponents; see e.g. [7,17,28].

An example that embodies these challenges is the family of 1D maps $f_a : [-1, 1] \circlearrowleft$ given by $f_a(x) = 1 - ax^2$, $a \in [0, 2]$. For this family, it has been proved that there is an open and dense set of parameters $\mathcal{A}$ such that for $a \in \mathcal{A}$, the orbit of Lebesgue-a.e. x tends to a sink under f_a [12,18]; at the same time there is a positive measure set $\mathcal{B}$ of parameters such that for $a \in \mathcal{B}$, f_a has an invariant density with respect to which it has a positive Lyapunov exponent [14] and exponential decay of time correlations [27]. Moreover, it has been shown that $\mathcal{A}$ and $\mathcal{B}$ together comprise a set of full Lebesgue measure in parameter space [19]. Given the complicated geometric relationship between $\mathcal{A}$ and $\mathcal{B}$, for a randomly picked parameter a, then, to determine for certain that it is in $\mathcal{B}$ (and not in $\mathcal{A}$) will require that we know the infinite-time dynamics of f_a to infinite precision.

Though less complete, the dynamical picture above has been shown to carry over to 2D, as exemplified in Newhouse's infinitely many sinks [21] and "strange attractors" for Hénon-like families [2,3,24]. See also [10,11,25,26]. These results suggest that away from systems with certain uniform properties, the dynamical landscape can be complex, with different regimes—some chaotic some not—coexisting in close proximity. In such situations, elaborate schemes of parameter selection are needed to identify maps with chaotic behavior.

For more references on the positivity of Lyapunov exponents, see the Introduction of [4].

Random maps

For randomly perturbed dynamical systems, especially those for which transition probabilities have densities, there are no issues with invariant measures. One can also hope that through the averaging effects of randomization, Lyapunov exponents become more stable and therefore easier to control. This and other ideas have been used a number of times in the past to prove continuity properties and simplicity of Lyapunov exponents in the settings of (i) random compositions of matrices and (ii) linear skew products over hyperbolic and partially hyperbolic systems; see [4] for references.

A stronger conjecture is that the Lyapunov exponents of a randomly perturbed system with expansion should more closely reflect the amount of expansion in the maps (in addition to being just positive), and a similar statement should hold for statistical quantities such as rate of mixing. General results of this type remain to be proved. The present paper is a step in that direction; two related results are [4,16].

Results of this paper

The maps considered here are dissipative versions of those in [4]. We consider small random perturbations of a class of (dissipative) deterministic 2D maps which have large regions $\mathcal{G}$ in their phase spaces with the property that the map is strongly uniformly hyperbolic for as long as an orbit remains in this region. The sets $\mathcal{G}$ are not invariant, however, and when an orbit returns to $\mathcal{G}$ after an excursion, its previously stable and unstable directions can be switched, resulting in cancellation of the type discussed above. Two results are proved: In Theorems 1 and 3, we prove the positivity of Lyapunov exponents with desirable bounds, and in Theorems 2 and 4, we prove the exponential decay of correlations, connecting the rate of decay to the system's Lyapunov exponents.

A salient difference between area-preserving and dissipative maps is that the latter can develop sinks, which even in the random setting are clear obstructions to both positive Lyapunov exponents and fast mixing. Elliptic islands in conservative systems slow down derivative growth as well, but the obstructions they impose are less severe because orbits can move away more easily. Another difference is that one has no *a priori* control on the statistics of random compositions of dissipative maps, whereas the statistics of random compositions of volume preserving maps are, tautologically, governed by the preserved volume.

Two simple facts used in our proof of positive Lyapunov exponents are (i) we assume perturbation size is large enough to ensure escape from the bad region, so that most orbits spend a large fraction of time in $\mathcal{G}$, and (ii) our randomization affects not only positions of x but also vector angles, so that cancellations of the type discussed above occur only with low probability. Less obvious is the crucial observation that one can achieve this control in a finite number of steps, and that the result can be shown to be uniform among initial locations x and vectors v. This is what enabled us to deduce, from *finite-time dynamics* alone, a lower bound for Lyapunov exponents as we have done in Theorems 1 and 3. For the reasons explained above, the proofs of these results are quite different from those in [4].

While noise alone can cause exponential decay in random dynamical systems: such decay rates are guaranteed only to be commensurate with the size of the noise. We prove in the setting above that for random perturbations of size $O(L^{-1})$, where L is the expansion constant, correlation decays at rate $O(L^{-\sigma})$ for some $\sigma > 0$. As we will show, once an initial condition is sufficiently randomized by the noise to acquire a density on a local scale, expanding properties of the map will, in a finite number of steps, spread this density from a local to global scale. Coupling time is related to the number of steps this takes, and that in turn is related to the Lyapunov exponent of the random maps system.

Finally, though we have chosen to illustrate our ideas in the simpler setting of 2D maps, our methods are not inherently dimension dependent, and we expect that similar ideas are valid in higher dimensions.

1. Setting and Statement of Results

Deterministic maps from which to perturb

This paper is about small random perturbations of a class of deterministic 2D-maps which we now describe. Let $\mathbb{S}^1 = \mathbb{R}/\mathbb{Z}$ be the unit circle. We assume throughout that $\psi : \mathbb{S}^1 \to \mathbb{R}$ is a C^3 function for which the following generic conditions hold:

(H1) $C'_\psi = \{\hat{x} \in \mathbb{S}^1 : \psi'(\hat{x}) = 0\}$ and $C''_\psi = \{\hat{z} \in \mathbb{S}^1 : \psi''(\hat{z}) = 0\}$ have finite cardinality.

(H2) $\min_{\hat{x} \in C'_\psi} |\psi''(\hat{x})| > 0$ and $\min_{\hat{z} \in C''_\psi} |\psi'''(\hat{z})| > 0$.

Abusing notation slightly, we will identify $\mathbb{S}^1$ with the half-open interval [0, 1), and write "x mod 1" for $\pi(x)$, where $\pi : \mathbb{R} \to \mathbb{S}^1$ is the usual projection.

For $L > 1$ and $a \in [0, 1)$, we define $f = f_{L,a} : \mathbb{S}^1 \to \mathbb{R}$ by $f(x) = L\psi(x) + a$. The deterministic map to be perturbed is $F = F_{L,a,b} : \mathcal{C} \to \mathcal{C}$, where $\mathcal{C} := \mathbb{S}^1 \times [0, b]$, $b \in (0, 1]$, and

$$F(x, y) = \begin{pmatrix} f(x) - y \mod 1 \\ bx \end{pmatrix}. \tag{1}$$

Notice that $\det(dF) \equiv b$. For $b = 1$, F can be seen as a diffeomorphism of the 2-torus $\mathbb{S}^1 \times \mathbb{S}^1$ onto itself. For $b < 1$, F is discontinuous at $\mathcal{D} = \{x = 0\}$; it can be seen as an embedding $\bar{F}$ of the rectangle $[0, 1] \times [0, b]$ into $\mathbb{S}^1 \times [0, b]$ with $\bar{F}(\{x = 0\}) \subset \{y = 0\}$ and $\bar{F}(\{x = 1\}) \subset \{y = b\}$.

For simplicity we will assume throughout that $0 \notin C'_\psi$.

Some examples of much studied systems that have a similar flavor to the setup above are (i) the Hénon maps [2,3,13], which correspond to the case $\psi(x) = x(1 - x)$ and $a = 0$, (ii) annulus maps such as those arising from kicked oscillators [25,26], with $\psi(x) = \sin(2\pi x)$, and (iii) the standard map [4,8], with $b = 1$ and $\psi(x) = \sin(2\pi x)$.

Random perturbations

We consider random compositions

$$F^n_{\underline{\omega}} = F_{\omega_n} \circ \cdots \circ F_{\omega_1} \qquad \text{for} \quad n = 1, 2, \ldots,$$

where $F_\omega = F \circ S_\omega$ and S_ω is of the form

$$S_\omega(x, y) = (x + \omega \bmod 1, y).$$

Here $\underline{\omega} = (\omega_1, \omega_2, \ldots)$, where $\omega_i \in [-\epsilon, \epsilon]$ are chosen *i.i.d.* with respect to the uniform distribution ν^ϵ on $[-\epsilon, \epsilon]$. Thus our sample space can be written as $\Omega = [-\epsilon, \epsilon]^{\mathbb{N}}$, equipped with the probability $\left(\nu^\epsilon\right)^{\otimes\mathbb{N}}$.

The random maps system above $\{F^n_{\underline{\omega}}\}_{n\geq 1}$ can also be seen as a time-homogeneous Markov chain $\{(X_n, Y_n)\}$ given by

$$(X_n, Y_n) = F^n_{\underline{\omega}}(X_0, Y_0) = F_{\omega_n}(X_{n-1}, Y_{n-1}).$$

That is to say, for fixed ϵ, the transition probability starting from $(x, y) \in \mathcal{C}$ is

$$P((x, y), A) = P^\epsilon((x, y), A) = \nu^\epsilon\{\omega \in [-\epsilon, \epsilon] : F_\omega(x, y) \in A\}$$

for Borel sets $A \subset \mathcal{C}$. For $k \geq 1$, we write P^k for the corresponding k-step transition probability, and for an initial distribution μ, μP^k is the distribution given by

$$\mu P^k(A) = \int P^k((x, y), A)d\mu((x, y)).$$

Results

Let "Leb" denote Lebesgue measure on $\mathcal{C}$. We view $b \in (0, 1]$ and ψ as fixed, and allow implicitly all constants to depend on them.

Theorem 1. *Given $\delta > 0$ and $\alpha < 1$, there exists $L_1 = L_1(\delta, \alpha)$, $L_1 \to \infty$ as $\delta \to 0$ or $\alpha \to 1$, such that the following holds for all $a \in [0, 1)$, $L \geq L_1$ and $\epsilon \geq L^{-1+\delta}$: For every $(x, y) \in \mathcal{C}$,*

$$\lambda_1^\epsilon := \lim_{n\to\infty} \frac{1}{n} \log \| \left(dF^n_{\underline{\omega}}\right)_{(x,y)} \| \geq \alpha \log L \qquad \text{for a.e. } \underline{\omega}. \tag{2}$$

In particular, the limit exists and is constant for all the (x, y) and $\underline{\omega}$ above.

As $\|dF\| = O(L)$, the lower bound on λ_1^ϵ in Theorem 1 is effectively as large as can be. This result requires no parameter deletion, i.e., it applies to all a, in sharp contrast to the situation for deterministic maps of a similar type (as in [2,3,14,26]). Fixing $b \in (0, 1)$ and considering all sufficiently large L, our minimum perturbation size ϵ is, at this level of generality, optimal; see the discussion in Sect. 2.1 C.

Our next result is about the rate of correlation decay for the Markov chain $\{(X_n, Y_n)\}$.

Theorem 2. *Given $\delta \in (0, 1)$, there exist constants $\sigma_1 = \sigma_1(\delta) > 0$, $K_1 = K_1(\delta) \in \mathbb{Z}^+$ and $L_2 = L_2(\delta) > 1$ such that the following holds for all $a \in [0, 1)$, $L \geq L_2$ and $\epsilon \geq L^{-1+\delta}$: If $\phi : \mathcal{C} \to \mathbb{R}$ is a bounded measurable function and μ_1, μ_2 are Borel probability measures on $\mathcal{C}$, then*

$$\left| \int \phi \, d(\mu_1 P^n) - \int \phi \, d(\mu_2 P^n) \right| \leq \|\phi\|_\infty \, L^{-\sigma_1 (n-K_1)} \qquad \text{for all } n \geq K_1.$$

Next we assume some minimal, easily checkable, condition on the first iterates of F, and show that the results above continue to hold for a significantly smaller ϵ. For $c > 0$, let $\mathcal{N}_c(C'_\psi)$ denote the c-neighborhood of C'_ψ in $\mathbb{S}^1$.

(H3)$_c$ For any $\hat{x}, \hat{x}' \in C'_\psi$, we have that $f\hat{x} - b\hat{x}' \pmod 1 \notin \mathcal{N}_c(C'_\psi)$.

The meaning of this condition is explained in Sect. 2.1. Theorems 3 and 4 are analogs of Theorems 1 and 2.

Theorem 3. *Fix arbitrary $c_0 > 0$. Then given $\delta > 0$ and $\alpha < 1$, there exists $L_3 = L_3(c_0, \delta)$ such that for all $L \geq L_3$ and $\epsilon \geq L^{-2+\delta}$, the conclusion of Theorem 1 holds for all $a \in [0, 1)$ for which $f = f_{L,a}$ satisfies (H3)$_{c_0}$.*

Theorem 4. *Fix arbitrary $c_0 > 0$. Given $\delta \in (0, 1)$, there are constants $\sigma_2 = \sigma_2(\delta) > 0$, $K_2 = K_2(\delta) \in \mathbb{Z}^+$ and $L_4 = L_4(c_0, \delta)$ such that for all $L \geq L_4$, $\epsilon \geq L^{-2+\delta}$ and a with $f = f_{L,a}$ satisfying (H3)$_{c_0}$, the conclusion of Theorem 2 holds with σ_2 and K_2 in the place of σ_1 and K_1 respectively.*

Not surprisingly, the results above are simpler for larger ϵ, though for technical reasons some of the proofs have to be written differently. We assume in all of the proofs below that $\epsilon \leq \frac{1}{2}$; the case of $\epsilon > \frac{1}{2}$ is left to the reader.

2. Preliminaries

The random maps we study are small perturbations of deterministic maps F of a particular form. In Sect. 2.1, we identify some geometric properties of F that will be relevant. We will also explain the relation between ϵ, the minimum perturbation size we require, and L, the approximate norm of dF. Stationary measures of the Markov chain $\{(X_n, Y_n)\}$ are discussed in Sect. 2.2.

2.1. Geometry of the deterministic map F.

A. Hyperbolicity The most salient characteristics of F are that (a) it is strongly hyperbolic on an open set that occupies a large fraction of the phase space and (b) this set is not invariant, and the hyperbolicity cannot be extended to all of $\mathcal{C}$. For F as defined in Sect. 1, consider cone fields in tangent spaces of the form $\mathcal{S} = \{v = (v_x, v_y) : |v_y/v_x| \leq \frac{1}{5}\}$. For $(x, y) \notin \{|f'| < 10\}$, one checks easily that $dF_{(x,y)}$ maps $\mathcal{S}$ into $\mathcal{S}$, and expands vectors in these cones uniformly. By (H1) and (H2), the region $\{|f'| < 10\}$ is comprised of a finite number of vertical strips in $\mathcal{C}$, the widths of which tend to 0 as $L \to \infty$, yet the cone preservation property above cannot be extended across these strips. Indeed it is not hard to convince oneself that there can be no continuous family of cones preserved by dF on all of $\mathcal{C}$, so F is not uniformly hyperbolic.

Equally important is how hyperbolicity deteriorates as one approaches the "bad set" $C'_\psi \times [0, b]$. Let $D_1 > 0$ be such that

$$|\psi'(x)| \geq 2D_1^{-1} d(x, C'_\psi)\,;$$

that such a D_1 exists follows from (H1), (H2) in Sect. 1. The following easy estimate is used many times:

Lemma 5. *For any $\eta \in (0, 1)$ and any $L > 1$,*

$$d\left(x, C'_\psi\right) > D_1 L^{-1+\eta} \implies |f'(x)| \geq 2L^\eta.$$

Lemma 5 has motivated us to work with sets of the form

$$B_\eta = \left\{(x, y) \in \mathcal{C} : d\left(x, C'_\psi\right) \leq D_1 L^{-1+\eta}\right\} \tag{3}$$

outside of which the dynamics are hyperbolic.

Finally, we observe that because we assumed that $0 \notin C'_\psi$ in Sect. 1, we have that the discontinuity line $\mathcal{D}$ is disjoint from $C'_\psi \times [0, b]$. Although it is perhaps technically unnecessary to assume this, we do so for the sake of clarity in our arguments.

B. Growth of horizontal curves Let us call a connected C^1 curve $\gamma \subset \mathcal{C}$ a *horizontal curve* if locally it is the graph of a map $h : I \to [0, b]$ where $I \subset \mathbb{S}^1$ is an interval and h satisfies $|h'| \leq 1/5$. We define the "length" of γ, denoted $|\gamma|$, to be the length of its projection to $\mathbb{S}^1$ counted with multiplicity, and say γ *wraps around* $\mathcal{C}$ if $|\gamma| \geq 1$. It follows immediately from Paragraph A that if γ is a horizontal curve defined on $I \subset \{|f'| \geq 10\} \setminus \mathcal{D}$, then $F(\gamma)$ is again a horizontal curve, and if γ is a horizontal curve disjoint from $B_\eta \cup \mathcal{D}$, $\eta \in (0, 1)$, then by Lemma 5,

$$|F(\gamma)| \geq L^\eta \cdot |\gamma|. \tag{4}$$

The next lemma gives some flavor of how horizontal curves grow in length until a component wraps around $\mathcal{C}$. More refined versions of this result will be proved later.

Lemma 6. *Given $\delta \in (0, 1)$, there exist $L_0 = L_0(\delta)$ such that for any $L \geq L_0$ and any horizontal curve γ with $|\gamma| \geq L^{-1+\delta}$, $F^n(\gamma)$ contains a horizontal curve that wraps around $\mathcal{C}$ for all $n \geq \frac{\log(4/\delta)}{\log 2}$.*

Proof. Assuming L is large enough, we observe first that if $|\gamma| > L^{-49/100}$, then $F^n(\gamma)$ contains a horizontal curve which wraps around $\mathcal{C}$ for all $n \geq 1$. This is because $\gamma \backslash (B_{1/2} \cup \mathcal{D})$ has a component $\hat{\gamma}$ of length $\geq \frac{1}{3} L^{-49/100}$, so by (4), $|F(\hat{\gamma})| > L^{1/2} \cdot \frac{1}{3} L^{-49/100} \gg 2$.

Suppose now that $|\gamma| < L^{-49/100}$. Observe that $\gamma \not\subset B_{\frac{3}{4}\delta}$ since $|\gamma| \geq L^{-1+\delta} \gg 2 D_1 L^{-1+\frac{3}{4}\delta}$, and $\gamma \backslash (B_{\frac{3}{4}\delta} \cup \mathcal{D})$ has at most two connected components since $|\gamma| < L^{-49/100}$. Let $\hat{\gamma}$ be the longer of these components. If $F(\hat{\gamma})$ wraps around $\mathcal{C}$, we are done. If not, we let $\gamma_1 = F(\hat{\gamma})$, and repeat the argument above with γ_1 in the place of γ.

More precisely, the longer component $\hat{\gamma}$ has length $|\hat{\gamma}| \geq \frac{1}{2}(L^{-1+\delta} - 2 D_1 L^{-1+3\delta/4}) \geq \frac{1}{3} L^{-1+\delta}$, and so by (4), $|\gamma_1| > \frac{1}{3} L^{-1+\frac{7}{4}\delta} > L^{-1+\frac{3}{2}\delta}$. If $|\gamma_1| > L^{-49/100}$, then we finish. If not, we let $\hat{\gamma}_1$ be the longer component of $\gamma_1 \backslash (B_{\frac{5}{4}\delta} \cup \mathcal{D})$, and conclude that $F(\hat{\gamma}_1) = \gamma_2$ has length $\geq \frac{1}{3} L^{-1+(\frac{3}{2}+\frac{5}{4})\delta} \geq L^{-1+\frac{5}{2}\delta}$.

Inductively, assume that we have produced curves $\gamma_3, \gamma_4, \ldots, \gamma_k$ following the above procedure, and that $L^{-1+\delta_k} \leq |\gamma_k| \leq L^{-49/100}$, where $\delta_k := \frac{1}{2}(1+2^k)\delta$. We set $\hat{\gamma}_k$ to be the longer of the two components of $\gamma_k \backslash (\mathcal{D} \cup B_{\delta_k - \frac{1}{4}\delta})$. Noting that $|\hat{\gamma}_k| \geq \frac{1}{3} L^{-1+\delta_k}$, it follows by (4) that $|\gamma_{k+1}| = |F(\hat{\gamma}_k)| \geq L^{\delta_k - \delta/4} \cdot \frac{1}{3} L^{-1+\delta_k} \geq \frac{1}{3} L^{-1+\frac{\delta}{4}+\delta_{k+1}} \geq L^{-1+\delta_{k+1}}$, since $\delta_{k+1} = 2\delta_k - \delta/2$.

In particular, we are ensured that $|\gamma_k| \geq L^{-49/100}$ if $k \geq \frac{\log(2/\delta)}{\log 2}$, and may now appeal to the argument for $|\gamma| \geq L^{-49/100}$ as above. □

C. Sinks and minimum sizes of perturbations Since we have not imposed conditions on F to rule out the possibility of a sink, to obtain positive Lyapunov exponents, our perturbation must be large enough to enable orbits to leave the sink's basin. Below we consider the worst possible case, when F has a sink $(\hat{x}, \hat{y})$ on $C'_\psi \times [0, b]$, and determine the size of the immediate basin for such a sink.

All norms in the next lemma refer to the max norm, i.e.

$$d_{\max}((x_1, y_1), (x_2, y_2)) := \max\{|x_1 - x_2|, |y_1 - y_2|\}.$$

Let $\mathfrak{B}((x, y), r) = \{(x', y') \in \mathcal{C} : d_{\max}((x, y), (x', y')) < r\}$.

Lemma 7. *Assume that $b < 1$. Let L be sufficiently large, and assume that $F(\hat{x}, \hat{y}) = (\hat{x}, \hat{y})$ with $\hat{x} \in C'_\psi$. Then,*

$$d_{\max}\big(F^2(x, y), (\hat{x}, \hat{y})\big) \leq \frac{1+b}{2}\, d_{\max}\big((x, y), (\hat{x}, \hat{y})\big) \qquad \textit{for all } (x, y) \in \mathfrak{B}(r_L)\ ,$$

where $\mathfrak{B}(\cdot) = \mathfrak{B}((\hat{x}, \hat{y}), \cdot)$ and $r_L := (L\|\psi''\|_{C^0})^{-1} \cdot \min\{\frac{1}{10}, \frac{1-b}{6}\}$.

Proof. As $0 \notin C'_\psi$, we may assume L is sufficiently large that $\mathfrak{B}(2r_L) \cap \mathcal{D} = \emptyset$.

Observe that if $(x, y) \in \mathfrak{B}(r_L)$, then $|f'(x)| \leq L\|\psi''\|_{C^0}|x - \hat{x}| \leq \kappa$ where $\kappa = \min\{\frac{1}{10}, \frac{1-b}{6}\}$. As one can check, for such (x, y) we have

$$\|dF_{(x,y)}\|_{\max} \leq \max\{|f'(x)| + 1, b\} \leq 1 + \kappa\ ,$$

guaranteeing $F(\mathfrak{B}(r_L)) \subset \mathfrak{B}((1+\kappa)r_L) \subset \mathfrak{B}(2r_L)$. Here, $\|\cdot\|_{\max}$ refers to the matrix norm induced by the max norm on $\mathbb{R}^2$.

We now estimate $\|dF^2_{(x,y)}\|_{\max}$ for $(x, y) \in \mathfrak{B}(r_L)$. Let $(x', y') = F(x, y)$. Then

$$dF^2_{(x,y)} = \begin{pmatrix} f'(x') & -1 \\ b & 0 \end{pmatrix} \begin{pmatrix} f'(x) & -1 \\ b & 0 \end{pmatrix} = \begin{pmatrix} f'(x)f'(x') - b & -f'(x') \\ bf'(x) & -b \end{pmatrix},$$

from which we obtain

$$\begin{aligned} \|dF^2_{(x,y)}\|_{\max} &\leq \max\{|f'(x)f'(x')| + b + |f'(x')|, b|f'(x)| + b\} \\ &\leq 2\kappa^2 + 2\kappa + b\,, \end{aligned}$$

having used $|f'(x')| \leq 2\kappa$. Since $\kappa \leq 1/10$, we have $2\kappa^2 \leq \kappa$, and so $\|dF^2_{(x,y)}\|_{\max} \leq 3\kappa + b \leq \frac{1}{2}(1 + b)$. The desired contraction estimate follows. □

Thus without knowing detailed properties of F, a difference between horizontal curves of length $O(L^{-1})$ and those of length $\gtrsim L^{-1+\delta}$, $\delta > 0$, is that the latter will always grow long under iterations of F (Lemma 6) whereas the first may not. This is a motivation for the size of ϵ in Theorem 1, which is, in the above sense, sharp. On the other hand, having a positive probability of escaping from a sink is a necessary but not sufficient condition. Whether or not $\lambda_1^\epsilon \approx \log L$ depends strongly on the asymptotic distribution of mass.

We remark that the sharpness of our lower estimate on ϵ is in the sense of fixing b and asking for a minimum perturbation size $\epsilon = \epsilon(L)$ for which the results hold for all large enough L. If we covary b and L, the picture may be quite different. For example, if $b = b(L) \to 1$ sufficiently fast as $L \to \infty$, then the ϵ needed is likely much smaller, as the results of [4] suggest.

D. Sizes of perturbations in Theorems 3 and 4 If the basins of sinks give a hint on the size of perturbation needed to achieve positivity of Lyapunov exponents, one may expect that ϵ can be shrunk if one requires, for example, that $F(C'_\psi \times [0, b])$ stays away from itself. There is validity to this idea, except that the condition cannot always be met for b large: the image of a vertical line $\{x\} \times [0, b]$ is a horizontal line of length b. In Theorems 3 and 4, we assert that in the case of random maps, to reduce the size of ϵ it suffices to impose the condition that no orbit can stay close to $C'_\psi \times [0, b]$ for *three consecutive iterates*. That is to say, suppose $F(x_i, y_i) = (x_{i+1}, y_{i+1}),\ i = 1, 2, \ldots$. If $x_i, x_{i+1} \in C'_\psi$, then x_{i+2} must stay away from C'_ψ. This is the meaning of (H3). Such a condition is both realizable and checkable, as it involves only a finite number of iterates for a finite set of points.

Theorems 3 and 4 suggest that it may be possible to decrease ϵ further if one imposes conditions on more iterates of f or F. We will not pursue that here, but this line of thinking is consistent with earlier ideas used in, e.g., [1,2,14,26] for proving nonuniform hyperbolicity for certain *deterministic* maps.

2.2. Stationary distributions. For the Markov chain $\{(X_n, Y_n)\}$ introduced in Sect. 1, a Borel probability measure μ on $\mathcal{C}$ is called *stationary* if for any Borel set $A \subset \mathcal{C}$,

$$\mu(A) = \int_{\mathcal{C}} P((x, y), A)\, d\mu(x, y).$$

Lemma 8. *For any $L, \epsilon > 0$, there exists a stationary probability measure for the Markov chain $\{(X_n, Y_n)\}$. Any such stationary probability is absolutely continuous with respect to Lebesgue measure.*

Proof. Since $\mathbb{S}^1 \times [0, b]$ is compact, to prove the existence stationary probability measures it suffices to check that $P((x, y), \cdot)$ varies continuously with $(x, y) \in \mathcal{C}$ in the weak* topology of Borel probability measures on $\mathcal{C}$. That is easily done.

That any stationary measure is absolutely continuous follows from the fact that $P^2((x, y), \cdot) \ll \text{Leb}$ for every $(x, y) \in \mathcal{C}$. One can check this directly, but given the special form of our map F a quick proof is that

$$F_{\omega_2} \circ F_{\omega_1}(x, y) = F \circ F' \circ S'_{\omega_1, -\omega_2}(x, y)\,, \tag{5}$$

where $S'_{\omega,\omega'}(x, y) = \big(x + \omega \bmod 1, y + \omega'\big)$ and F' is the extension of F in (1) to $\mathbb{S}^1 \times [-\epsilon, b + \epsilon]$. $\square$

Given that the densities of all stationary distributions are supported on sets with Lebesgue measure greater than constant$\cdot\epsilon^2$, there can be at most a finite number of ergodic probability measures. From the Multiplicative Ergodic Theorem for random systems [15], we have that with respect to any stationary measure, the limit in λ_1^ϵ as defined in (2) exists a.e. and is constant on each ergodic component.

Theorems 1 and 2 will be proved as follows: We will show in Sect. 3 that for any stationary measure μ, $\lambda_1^\epsilon \geq \alpha \log L$ for μ-a.e. (x, y). Theorem 2 is proved in Sect. 4. From Theorem 2, it follows easily that (i) there can be at most one stationary measure μ, and (ii) for *every* $(X_0, Y_0) \in \mathcal{C}$,

$$\mathbb{P}_{(X_0,Y_0)}[(X_n, Y_n) \in \text{supp}(\mu) \text{ for some } n] = 1. \tag{6}$$

Lemma 8, (i) and (ii) above, together with the Lyapunov exponent estimate in Sect. 3 complete the proof of Theorem 1. Theorems 3 and 4 are proved analogously; the proofs are given in Sect. 5.

3. Estimation of Lyapunov Exponents

In this section we consider an arbitrary stationary measure μ and prove that $\lambda_1^\epsilon \geq \alpha \log L$ for μ-a.e. (x, y). We will refer to this result as the "μ-a.e. version Theorem 1". As with ψ and b, we assume a is fixed throughout. For technical reasons, we consider only $\delta \in (0, 1)$, and let $K = K(\delta)$ be the smallest integer for which $2^K \delta > 4$. This number has the connotation of a renewal time and will appear in many of our proofs. Notice that it is independent of L.

3.1. Reductions. Below, we use the notation $u_\theta = (\cos\theta, \sin\theta)$ for $\theta \in [0, 2\pi)$. Given an initial condition $(X_0, Y_0) \in \mathcal{C}$, an "angle" $\theta_0 \in [0, 2\pi)$, and a sequence of perturbations $\omega_1, \ldots, \omega_n \in [-\epsilon, \epsilon]$, we write $(X_k, Y_k) = F_{\omega_k}(X_{k-1}, Y_{k-1}),\ k = 1, \ldots, n$, and let $\theta_k \in [0, 2\pi)$ be defined by

$$u_{\theta_k} = \frac{(dF_{\omega_k})_{(X_{k-1},Y_{k-1})} u_{\theta_{k-1}}}{\|(dF_{\omega_k})_{(X_{k-1},Y_{k-1})} u_{\theta_{k-1}}\|}.$$

Our first reduction is to an integral that involves only *finite-time* dynamics.

Proposition 9. *Given $\alpha < 1$ and $\delta \in (0, 1)$, the following holds for all L sufficiently large and all $\epsilon \in [L^{-1+\delta}, \frac{1}{2}]$: for any $(X_0, Y_0) \in \mathcal{C}$ and $\theta_0 \in [0, 2\pi)$, we have the uniform bound*

$$\int \log \|(dF_{\omega_{K+3}})_{(X_{K+2},Y_{K+2})} u_{\theta_{K+2}}\| \, (d\nu^\epsilon)^{\otimes(K+3)}(\omega_1, \ldots, \omega_{K+3}) \geq \alpha \log L\,, \qquad (7)$$

where K is as defined at the beginning of this section.

The idea behind Proposition 9 is that starting from any (X_0, Y_0) and θ_0, after $K+2$ steps, the distributions of (X_{K+2}, Y_{K+2}) and θ_{K+2} are sufficiently well distributed that the average growth of $u_{\theta_{K+2}}$ under $dF_{\omega_{K+3}}$ is large.

Proof of μ-a.e. version of Theorem 1 assuming Proposition 9. For μ-a.e. (X_0, Y_0), the convergence in (2) holds for a.e. $\underline{\omega}$. It therefore converges in L^1 by the Bounded Convergence Theorem. That is,

$$\lambda_1^\epsilon = \lim_{n\to\infty} \frac{1}{n} \int_\Omega \log \|(dF^n_{\underline{\omega}})_{(X_0,Y_0)}\| d\mathbb{P}(\underline{\omega}).$$

Fix an arbitrary $\theta_0 \in [0, 2\pi)$, and let $\{\theta_n\}$ be as defined above. Then

$$\begin{aligned}\lambda_1^\epsilon &= \lim_{n\to\infty} \frac{1}{n} \int_\Omega \log \|(dF^n_{\underline{\omega}})_{(X_0,Y_0)}\| d\mathbb{P}(\underline{\omega}) \geq \lim_{n\to\infty} \frac{1}{n} \int_\Omega \log \|(dF^n_{\underline{\omega}})_{(X_0,Y_0)} u_{\theta_0}\| d\mathbb{P}(\underline{\omega}) \\ &= \lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \int_\Omega \log \|(dF_{\omega_{i+1}})_{(X_i,Y_i)} u_{\theta_i}\| d\mathbb{P}(\underline{\omega}). \qquad (8)\end{aligned}$$

For $n \gg K$ and $K+2 \leq i \leq n-1$, the i-th summand above is equal to

$$\int \left(\int \log \|(dF_{\omega_{i+1}})_{(X_i,Y_i)} u_{\theta_i}\| (d\nu^\epsilon)^{\otimes(K+3)}(\omega_{i-K-1}, \ldots, \omega_{i+1}) \right) (d\nu^\epsilon)^{\otimes(i-K-2)}(\omega_1, \ldots, \omega_{i-K-2}).$$

Observe that for each *fixed* $\omega_1, \ldots, \omega_{i-K-2}$, the parenthetical term has the same form as the LHS of (7) with the replacements $(X_0, Y_0) \mapsto (X_{i-K-2}, Y_{i-K-2})$ and $\theta_0 \mapsto \theta_{i-K-2}$. It is therefore $\geq \alpha \log L$ by Proposition 9, completing the proof of Theorem 1. □

A second reduction is given in Proposition 10 below. It asserts that to obtain the bound in Proposition 9, it is sufficient to leverage the randomness in only two time steps.

Proposition 10 (Main Proposition). *The assumptions are as in Proposition 9. Let $(X_0, Y_0) \in \mathcal{C}$ and $\theta_0 \in [0, 2\pi)$, and fix arbitrary $\omega_2, \ldots, \omega_K, \omega_{K+2}, \omega_{K+3} \in [-\epsilon, \epsilon]$. Then*

$$\int \log \|(dF_{\omega_{K+3}})_{(X_{K+2},Y_{K+2})} u_{\theta_{K+2}}\| \, d(\nu^\epsilon \otimes \nu^\epsilon)(\omega_1, \omega_{K+1}) \geq \alpha \log L. \qquad (9)$$

In Sect. 3.2, we will show that randomizing ω_1 alone will lead to favorable distributions of *mass* on $\mathcal{C}$. The control of *angles*, or slopes of tangent vectors, is the other crucial ingredient in the estimation of Lyapunov exponents for maps in dimensions greater than one. This control is achieved by varying both ω_1 and ω_{K+1}. The argument is carried out in Sect. 3.3, where the proof is completed.

3.2. Distribution of mass via a one-step randomization. The aim of this subsection is to show that by randomizing ω_1, the distribution of X_K given any (X_0, Y_0) and $\omega_2, \ldots, \omega_K$ has certain desirable characteristics.

The idea is as follows: Randomization of ω_1 produces a horizontal line segment, which we iterate forward using $F_{\omega_i} \circ \cdots \circ F_{\omega_2} \circ F_\omega$ for $i = 2, 3, \ldots, K$. As explained in Sect. 2.1A, on a large part of the phase space (which we will call "good"), cones of tangent vectors that are roughly horizontal are preserved and the lengths of the vectors are expanded by the derivatives of the maps. We will show that under our conditions on L and ϵ, a large fraction of the horizontal segment above will remain in this "good" region in the first K iterates and has a nice distribution at the end.

The idea of iterating densities on curves or submanifolds whose tangent vectors lie in unstable cones has been used many times in uniformly hyperbolic and other systems with globally defined invariant cones (including billiards); see e.g. [22,23], and also [9] where the concept of what is now called "standard pair" originated. Our situation is different in that our maps do not admit globally defined invariant cone families. In that respect our construction is closer to that in [2,26], where curves that are aligned with certain favorable conditions are pushed forward and those parts that are hard to control are deleted—the situation here being much simpler due to the large size of L and the fact that we need only to iterate a finite number of times. We mention also that the use of random perturbations to create segments or surfaces "in favorable directions" to be pushed forward was used in [20] to treat certain PDEs that are randomly forced in their low modes.

Let $I \subset \mathbb{S}^1$ be an interval. We say a horizontal curve in $\mathcal{C}$ *crosses* $I \times [0, b]$ if its graphing function is defined on I.

Proposition 11. *Given any $\delta \in (0, 1)$, there exist $L^* = L^*(\delta) \geq 1$ and $C = C(\delta) \geq 1$ such that the following hold for all $L \geq L^*$ and $\epsilon \in [L^{-1+\delta}, \frac{1}{2}]$: For every $(X_0, Y_0) \in \mathcal{C}$ and arbitrary $\omega_2, \ldots, \omega_K \in [-\epsilon, \epsilon]$, there exists a set $R \subset [-\epsilon, \epsilon]$ such that*

(a) $\nu^\epsilon([-\epsilon, \epsilon]\backslash R) \leq CL^{-\delta/4}$, *and*
(b) *on R there is a partition $\mathcal{W} = \{W\}$ into intervals with the following properties:*
(i) *the set $\{(X_K, Y_K) = F_{\omega_K} \circ \cdots \circ F_{\omega_2} \circ F_\omega(X_0, Y_0) : \omega \in W\}$ is a horizontal curve crossing $(0, 1) \times [0, b]$; and*
(ii) *viewing X_K as a function of $\omega \in W$, we have that for $\omega, \omega' \in W$,*

$$\left| \frac{\frac{\partial}{\partial \omega} X_K(\omega)}{\frac{\partial}{\partial \omega} X_K(\omega')} \right| \leq C. \tag{10}$$

To obtain distortion estimates, it will be necessary to control the second derivatives of horizontal curves. Let $\hat{D} = \|\psi''\|_{C^0}$, and let $\mathcal{H}$ denote the set of all connected C^2 curves γ in $\mathcal{C}$ that are finite unions of curves $\hat{\gamma}$ of the form $\hat{\gamma}$ =graph(h) where $h : I \to [0, b]$, $I \subset \mathbb{S}^1$ is an interval, $|h'| \leq \frac{1}{5}$ and $|h''| \leq \hat{D}L$.

Lemma 12. *Let $\gamma \in \mathcal{H}$ be such that $\gamma \subset \{|f'| \geq 10\}$ and $\gamma \cap \mathcal{D} = \emptyset$. Then $F(\gamma) \in \mathcal{H}$.*

This lemma will be used implicitly many times in the proofs below; its proof is straightforward and left to the reader.

In the proof of Proposition 11 and related proofs to follow, we will call a horizontal curve γ *short* if $|\gamma| \leq L^{-\frac{1}{2}+\delta/4}$, and *long* otherwise.

Proof of Proposition 11. Let $\delta \in (0, 1)$ be fixed; the magnitude of L may be increased a finite number of times as we go along. Fix arbitrary $(X_0, Y_0) \in \mathcal{C}$ and $\omega_2, \ldots, \omega_K \in [-\epsilon, \epsilon]$. We divide the proof into three parts.

(A) Inductive construction of R and $\mathcal{W}$. For $k = 0, 1, \ldots, K$, we define a collection Γ_k of horizontal curves inductively as follows. Let $\Gamma_0 = \{\gamma_0\}$, where $\gamma_0 := [X_0 - \epsilon, X_0 + \epsilon] \times \{Y_0\}$. We identify, via a procedure to be specified, a collection $\mathcal{R}_{\gamma_0}$ of subsegments of γ_0 and let $\Gamma_1 = \{F(\hat{\gamma}) : \hat{\gamma} \in \mathcal{R}_{\gamma_0}\}$. For $k \geq 1$, to obtain Γ_{k+1} from Γ_k, we identify for each $\gamma \in \Gamma_k$ a disjoint collection $\mathcal{R}_\gamma$ of subsegments $\hat{\gamma} \subset \gamma$, and let $\Gamma_{k+1} := \{F_{\omega_{k+1}}(\hat{\gamma}) : \hat{\gamma} \in \mathcal{R}_\gamma, \gamma \in \Gamma_k\}$. Observe that each $\gamma \in \Gamma_k$ so defined has the form

$$\gamma = \{F_{\omega_k} \circ \cdots \circ F_{\omega_2} \circ F_\omega(X_0, Y_0) : \omega \in W\}$$

for an interval $W \subset [-\epsilon, \epsilon]$. Moreover, if we let $\mathcal{W}_k$ denote the set of W associated with $\gamma \in \Gamma_k$, and let $R_k = \cup_{W \in \mathcal{W}_k} W$, then it follows from our construction that

$$[-\epsilon, \epsilon] =: R_0 \supset R_1 \supset \cdots \supset R_K.$$

The set R in Proposition 11 is R_K, and $\mathcal{W} = \mathcal{W}_K$.

To specify $\mathcal{R}_\gamma$ for $\gamma \in \Gamma_k$ involves deleting some parts of γ and partitioning the rest into subsegments. Our choices below are aimed at achieving the following: the fractions deleted must be small, and the partition must result in segments short enough for distortion control, yet expanded enough to reach full length in at most K iterations.

Consider first the case $\gamma = \gamma_0$:

Case 1. γ is *short*. We let $\delta_\gamma > 0$ be given by $|\gamma| = L^{-1+\delta_\gamma+\delta/4}$. First we delete $\gamma \cap (B_{\delta_\gamma} \cup \mathcal{D})$ (see Sect. 2.1 for the definition of B_η). Then we put into $\mathcal{R}_\gamma$ connected subsegments $\hat{\gamma}$ such that $\hat{\gamma} \cap B_{\frac{1}{2}} = \emptyset$ and $F(\hat{\gamma})$ crosses $(0, 1) \times [0, b]$. From what remains of γ, we discard those components that have length $< L^{-1+\delta_\gamma}$, partition the rest into connected subsegments of length between $L^{-1+\delta_\gamma}$ and $2L^{-1+\delta_\gamma}$ and put them into $\mathcal{R}_\gamma$.

Case 2. γ is *long*. We delete $\gamma \cap (B_{1/2} \cup \mathcal{D})$, and put into $\mathcal{R}_\gamma$ any connected subsegment $\hat{\gamma}$ that remains for which $F(\hat{\gamma})$ crosses $(0, 1) \times [0, b]$.

For $\gamma \in \Gamma_k, k \geq 1$, the procedure is a small modification of that above to accommodate the use of $F_{\omega_{k+1}}$ in the place of F. That is, when γ is short, we delete $\gamma \cap \big((B_{\delta_\gamma} \cup \mathcal{D}) - (\omega_{k+1}, 0)\big)$, and put into $\mathcal{R}_\gamma$ any connected subsegment $\hat{\gamma}$ that remains for which $F_{\omega_{k+1}}(\hat{\gamma})$ crosses $(0, 1) \times [0, b]$; the rest of the construction of Γ_{k+1} proceeds exactly as before. The case where γ is long is treated analogously.

This completes the construction.

In the rest of (A) we argue that all the curves in Γ_K cross $(0, 1) \times [0, b]$. This along with Lemma 12 will prove (b)(i) in Proposition 11. Observe that once $\gamma \in \Gamma_k$ is long, then all of its descendants, i.e. curves of the form $F_{\omega_{k+1}}(\hat{\gamma}), \hat{\gamma} \in \mathcal{R}_\gamma$, cross $(0, 1) \times [0, b]$, as will the descendants of $F_{\omega_{k+1}}(\hat{\gamma})$. Suppose $\gamma_0, \gamma_1, \gamma_2, \ldots$ are such that for all k, $\gamma_k \in \Gamma_k$, γ_{k+1} descends from γ_k, and all are short. We will show that δ_{γ_k} grows exponentially: First, since $L^{-1+\delta} \leq \epsilon = \frac{1}{2} L^{-1+\delta_{\gamma_0}+\delta/4}$, we may assume L is large enough that $\delta_{\gamma_0} \geq \frac{1}{2}\delta$. Suppose $\gamma_1 = F(\hat{\gamma})$. By construction, $|\hat{\gamma}| \geq L^{-1+\delta_{\gamma_0}}$, and by Lemma 5, $|f'| > 2L^{\delta_{\gamma_0}}$ on $\hat{\gamma}$. So $|\gamma_1| \geq |\hat{\gamma}| \cdot L^{\delta_{\gamma_0}} \geq L^{-1+2\delta_{\gamma_0}}$ by (4). As δ_{γ_1} is defined by $L^{-1+\delta_{\gamma_1}+\delta/4} = |\gamma_1| \geq L^{-1+2\delta_{\gamma_0}}$, it follows that $\delta_{\gamma_1} \geq 2\delta_{\gamma_0} - \frac{1}{4}\delta \geq \frac{3}{4}\delta$. Repeating this argument, we obtain inductively that

$$\delta_{\gamma_{k+1}} \geq 2\delta_{\gamma_k} - \frac{1}{4}\delta \geq \frac{1}{4}(2^{k+1}+1)\delta.$$

Since $2^K\delta > 4$ by our choice of K (beginning of Sect. 3), it follows that $L^{-1+\frac{1}{4}(2^{K-1}+1)\delta} > L^{-\frac{1}{2}+\delta/4}$, i.e., all $\gamma \in \Gamma_{K-1}$ are either long, or their F_{ω_K}-image crosses $(0,1)\times[0,b]$.

(B) Controlling distortion. Fix arbitrary $W \subset \mathcal{W}$. We let $\check{\gamma}_0 = \{(X_0+\omega, Y_0) : \omega \in W\}$, $\check{\gamma}_k = (F_{\omega_k} \circ \cdots \circ F_{\omega_2} \circ F)(\check{\gamma}_0)$ for $1 \leq k \leq K$, and let $\gamma_k \in \Gamma_k$ be such that $\check{\gamma}_k \subset \gamma_k$. We assume $\check{\gamma}_k = \mathrm{graph}(h_k)$ for $h_k : I_k \to [0,b]$, and let $\pi_1 : \mathcal{C} \to (0,1)$ be projection onto the first coordinate. Then defining

$$\begin{aligned} &\check{F}_1 : I_0 \to I_1 \quad \text{by} \quad \check{F}_1(x) = \pi_1 F(x, h_0(x)) \ , \\ \text{and} \quad &\check{F}_k : I_{k-1} \to I_k \quad \text{by} \quad \check{F}_k(x) = \pi_1 F_{\omega_k}(x, h_{k-1}(x)) \quad \text{for} \quad k > 1 \ , \end{aligned}$$

we have that for $\omega \in W$, $X_k(\omega) = \check{F}^k(X_0+\omega)$ where $\check{F}^k := \check{F}_k \circ \cdots \circ \check{F}_1$. We estimate the distortion of $\check{F}^k$ on I_0 as follows: For $x_1, x_2 \in I_0$,

$$\frac{(\check{F}^k)'(x_1)}{(\check{F}_k)'(x_2)} = \prod_{i=1}^{k} \frac{(\check{F}_i)'\left(\check{F}^{i-1}x_1\right)}{(\check{F}_i)'\left(\check{F}^{i-1}x_2\right)}. \tag{11}$$

If γ_{i-1} is long, then $\inf_{z\in I_{i-1}} |\check{F}_i'(z)| \geq L^{\frac{1}{2}}$ (by (4) and Lemmas 5, 12), while $|I_{i-1}|\cdot L^{\frac{1}{2}} \leq |I_i| \leq 1$, implying $|I_{i-1}| \leq L^{-\frac{1}{2}}$. Thus

$$\left|\log \frac{(\check{F}_i)'\left(\check{F}^{i-1}x_1\right)}{(\check{F}_i)'\left(\check{F}^{i-1}x_2\right)}\right| \leq \frac{\sup_{z\in I_{i-1}} |\check{F}_i''(z)|}{\inf_{z\in I_{i-1}} |\check{F}_i'(z)|} \cdot |I_{i-1}| \ \leq \ \frac{2\hat{D}L}{L^{\frac{1}{2}}} \cdot L^{-\frac{1}{2}} \ = \ 2\hat{D}. \tag{12}$$

If γ_{i-1} is short, the same estimate holds if $F_{\omega_i}(\gamma_{i-1})$ crosses $(0,1)\times[0,b]$. Otherwise

$$\begin{aligned} \left|\log \frac{(\check{F}_i)'(\check{F}^{i-1}x_1)}{(\check{F}_i)'(\check{F}^{i-1}x_2)}\right| &\leq \frac{\sup_{z\in I_{i-1}} |\check{F}_i''(z)|}{\inf_{z\in I_{i-1}} |\check{F}_i'(z)|} \cdot |I_{i-1}| \\ &\leq \frac{2\hat{D}L}{L^{\delta_{\gamma_{i-1}}}} \cdot 2L^{-1+\delta_{\gamma_{i-1}}} \ = \ 4\hat{D}. \end{aligned} \tag{13}$$

We have shown that the left side of (11) is $\leq e^{4k\hat{D}}$, completing the proof of (b)(ii) in Proposition 11.

(C) Estimating $\nu^\epsilon(R)$. Noting that K is independent of L, it suffices to estimate the fraction of mass deleted as we go from R_k to R_{k+1} for each $k < K$. To do that, we consider one $W \in \mathcal{W}_k$ at a time, beginning with γ_0. Let $\gamma \in \Gamma_k$ be associated with W. If γ is short, then for L large enough, γ can meet at most one of the following: $\mathcal{D}$, or a component B of $B_{\delta_\gamma} - (\omega_{k+1}, 0)$. Let $\bar{\gamma}$ be the union of those subsegments of γ whose images cross $(0,1)\times[0,b]$. Then $\gamma\backslash(\mathcal{D}\cup B\cup\bar{\gamma})$ has at most 4 connected components.

The total length deleted is $\le 2D_1L^{-1+\delta_\gamma} + 4L^{-1+\delta_\gamma}$. Since $|\gamma| = L^{-1+\delta_\gamma+\delta/4}$, pulling back to W and using the distortion bound in (B), we have

$$\frac{\nu^\epsilon\left(W\backslash \cup_{W'\in\mathcal{W}_{k+1}} W'\right)}{\nu^\epsilon(W)} \le e^{4k\hat{D}}(2D_1+4)L^{-\delta/4}.$$

If γ is long, i.e., $|\gamma| > L^{-\frac{1}{2}+\delta/4}$, then it can meet any number of the $\#C'_\psi$ components of $B := B_{\frac{1}{2}} - (\omega_{k+1}, 0)$, hence there are at most $\#C'_\psi + 1$ connected components γ' of $\gamma\backslash(\mathcal{D}\cup B)$. The $F_{\omega_{k+1}}$-image of each such γ' consists of a disjoint union of curves crossing $(0, 1)\times[0, b]$ and at most two additional curves of length < 1 to be deleted. If $\gamma'' \subset \gamma'$ is such that $F_{\omega_{k+1}}(\gamma'')$ is one of the curves of length < 1, then $|\gamma''| < L^{-\frac{1}{2}}$. Altogether, we obtain

$$\frac{\nu^\epsilon\left(W\backslash \cup_{W'\in\mathcal{W}_{k+1}} W'\right)}{\nu^\epsilon(W)} \le e^{4k\hat{D}}\left(1+\#C'_\psi\right)(2D_1+2)L^{-\delta/4}.$$

With L taken sufficiently large, the proof of part (a) of the Proposition is complete. □

3.3. Proof of main proposition (Proposition 10). We use the notation in Sect. 3.1, and assume the hypotheses of Proposition 10. In particular, throughout this subsection δ and α are fixed, as are (X_0, Y_0) and $\omega_2, \ldots, \omega_K, \omega_{K+2}, \omega_{K+3}$, leaving ω_1 and ω_{K+1} to be varied.

We introduce the following notation: For $c^* > 0$, let $B^1_{c^*} := \{x \in \mathbb{S}^1 : d(x, C'_\psi) \le D_1c^*\}$ where D_1 is as in Sect. 2.1, so that $|\psi'| \ge 2c^*$ outside of $B^1_{c^*}$. Similarly, by (H1) and (H2), there is a constant $D_2 > 0$ for which $|\psi''(x)| \ge D_2^{-1}d(x, C''_\psi)$ for all $x \in \mathbb{S}^1$. Let $B^2_{c^*} = \{x \in \mathbb{S}^1 : d(x, C''_\psi) \le D_2c^*\}$, so that outside of $B^2_{c^*}$, $|\psi''| \ge c^*$.

(A) Varying ω_1 (for fixed ω_{K+1}). Let $p > 0$ be small enough to be specified at the end of the proof, and let R be as in Proposition 11 with respect to the $\omega_2, \ldots, \omega_K$ specified. We assume L is large enough that $\nu^\epsilon(R) > 1 - p$. For each ω_{K+1}, we define

$$G_1(\omega_{K+1}) = \Big\{\omega_1 \in R \ : \ X_K + \omega_{K+1}, X_{K+1} + \omega_{K+2}, X_{K+2} + \omega_{K+3} \notin B^1_{c^*}, \\ \text{and } X_K + \omega_{K+1} \notin B^2_{c^*}\Big\},$$

and let

$$G_1 = \{(\omega_1, \omega_{K+1}) \in [-\epsilon, \epsilon]\times[-\epsilon, \epsilon] : \omega_1 \in G_1(\omega_{K+1})\}.$$

This is a “good set”, constructed to help us gain control of the integrand in Proposition 10.

We claim that for small enough $c^* = c^*(p)$ and large enough L depending on c^* and p, $\nu^\epsilon(G_1(\omega_{K+1})) > 1-2p$, the values of c^* and L required being independent of ω_{K+1}. Since $|\gamma| = 1$ for all $\gamma \in \Gamma_K$, both $F_{\omega_{K+1}}(\gamma)$ and $F_{\omega_{K+2}}F_{\omega_{K+1}}(\gamma)$ wrap around $\mathcal{C}$ many times, and so the portions of $\gamma + (\omega_{K+1}, 0)$, $F_{\omega_{K+1}}(\gamma) + (\omega_{K+2}, 0)$, $F_{\omega_{K+2}}F_{\omega_{K+1}}(\gamma) + (\omega_{K+3}, 0)$ that fall into $B^1_{c^*}$ tend to 0 as $c^* \to 0$. The same comment applies to the portion of $\gamma + (\omega_{K+1}, 0)$ that falls into $B^2_{c^*}$. Pulling back to ω_1 where the deletion takes place, we need a distortion bound not explicitly depending on c^*. This is possible as the quantity in (12) with $i = K+1$ or $K+2$ is $\leq (2\hat{D}L)/(Lc^*)^2 \ll 2\hat{D}$.

(B) Varying ω_{K+1} (for fixed ω_1). Our "good set" here is the Borel set

$$G_2 = \{(\omega_1, \omega_{K+1}) \in [-\epsilon, \epsilon] \times [-\epsilon, \epsilon] : \omega_{K+1} \in G_2(\omega_1)\},$$

$$\text{where} \qquad G_2(\omega_1) = \Big\{\omega_{K+1} \in [-\epsilon, \epsilon] : X_K + \omega_{K+1} \notin B^2_{c^*} \text{ and } |\tan\theta_{K+1}| \leq L^{1-\sigma}\Big\}.$$

Above, $\sigma > 0$ is a parameter to be specified shortly. As we will see, it is the condition on θ_{K+1} that is of interest; the one on $X_K + \omega_{K+1}$ is used in the estimation of θ_{K+1}.

Letting $G_2^c(\omega_1) = [-\epsilon, \epsilon] \backslash G_2(\omega_1)$, we write $G_2^c(\omega_1) = G_{2,1}^c(\omega_1) \cup G_{2,1}^c(\omega_1)$ where

$$G_{2,1}^c(\omega_1) = \Big\{\omega_{K+1} \in [-\epsilon, \epsilon] : X_K + \omega_{K+1} \in B^2_{c^*}\Big\}$$

$$\text{and} \qquad G_{2,2}^c(\omega_1) = \Big\{\omega_{K+1} \in [-\epsilon, \epsilon] : X_K + \omega_{K+1} \notin B^2_{c^*} \text{ and } |\tan\theta_{K+1}| \; L^{1-\sigma}\Big\}.$$

We do not estimate the measure of $G_{2,1}^c(\omega_1)$, as the pairs (ω_1, ω_{K+1}) with $\omega_{K+1} \in G_{2,1}^c(\omega_1)$ are already excluded in the definition of G_1. We will show that $\nu^\epsilon(G_{2,2}^c(\omega_1)) < p$ for large enough L independently of ω_1.

The tangent term in $G_{2,2}^c(\omega_1)$ is estimated using the formula

$$\tan\theta_{K+1} = \frac{b}{f'(X_K + \omega_{K+1}) - \tan\theta_K}, \tag{14}$$

which follows directly from the definition of dF. Rephrasing, we have that

$$\begin{aligned} G_{2,2}^c(\omega_1) &= \Big\{\omega \in [-\epsilon, \epsilon] : X_K + \omega \notin B^2_{c^*} \text{ and } |f'(X_K + \omega) \\ &\qquad - \tan\theta_K| < bL^{-1+\sigma}\Big\} \\ &= \Big\{\omega \in [-\epsilon, \epsilon] : X_K + \omega \notin B^2_{c^*} \text{ and } |\psi'(X_K + \omega) \\ &\qquad - \frac{1}{L}\big(\tan\theta_K - a\big)| < bL^{-2+\sigma}\Big\}. \end{aligned}$$

The set of ω for which $X_K + \omega \notin B^2_{c^*}$ consists of at most $\# C''_\psi$ intervals; on each interval, $\omega \mapsto \psi'(X_k + \omega)$ is monotonic with $|\psi''(X_k + \omega)| \geq c^*$. Thus independently of the value of $\frac{1}{L}\big(\tan\theta_K - a\big)$,

$$\nu^\epsilon\left(G_{2,2}^c(\omega_1)\right) < \# C''_\psi \cdot \frac{1}{2\epsilon} \frac{2bL^{-2+\sigma}}{c^*} \leq \# C''_\psi \cdot \frac{b}{c^*} L^{-1+\sigma-\delta}, \tag{15}$$

using $\epsilon \geq L^{-1+\delta}$. Taking $\sigma \leq 1$, this bound is $< p$ for L sufficiently large.

(C) Putting it all together. At last we form the "good set" $G = G_1 \cap G_2$, and note that $(\nu^\epsilon \otimes \nu^\epsilon)(G) > 1 - 3p$. To estimate (9), we decompose into $\int_G + \int_{G^c}$.

For $(\omega_1, \omega_{K+1}) \in G$, we use

$$\|(dF_{\omega_{K+3}})_{(X_{K+2}, Y_{K+2})} u_{\theta_{K+2}}\| \geq |\cos\theta_{K+2}| \cdot |f'(X_{K+2} + \omega_{K+3}) - \tan\theta_{K+2}|. \tag{16}$$

To bound the cosine factor, we have that

$$|\tan\theta_{K+2}| \leq \frac{b}{|f'(X_{K+1} + \omega_{K+2})| - |\tan\theta_{K+1}|} \leq \frac{b}{c^* L - L^{1-\sigma}}\,, \tag{17}$$

the second inequality following from the definitions of G_1 and G_2. A generous bound is therefore $|\cos\theta_{K+2}| \geq 1/2$. To estimate the second factor on the right side of (16), we have $|f'(X_{K+2} + \omega_{K+3})| \geq c^* L$ (property of G_1), and $|\tan\theta_{K+2}| = O(L^{-1})$ by (17). Altogether, this gives, for $(\omega_1, \omega_{K+1}) \in G$,

$$\|(dF_{\omega_{K+3}})_{(X_{K+2}, Y_{K+2})} u_{\theta_{K+2}}\| \geq \frac{c^*}{4} L.$$

For $(\omega_1, \omega_{K+1}) \in G^c$, we bound the integrand with the 'worst' possible estimate

$$\|(dF_{\omega_{K+3}})_{(X_{K+2}, Y_{K+2})} u_{\theta_{K+2}}\| \geq \frac{b}{2L\|\psi'\|_{C^0}}.$$

Collecting our estimates,

$$\begin{aligned}
&\left(\int_{(\omega_1, \omega_{K+1}) \in G} + \int_{(\omega_1, \omega_{K+1}) \in G^c}\right) \\
&\qquad \times \log \|(dF_{\omega_{K+3}})_{(X_{K+2}, Y_{K+2})} u_{\theta_{K+2}}\| d(\nu^\epsilon \otimes \nu^\epsilon)(\omega_1, \omega_{K+1}) \\
&\geq (1 - 3p) \log L + \log\left(\frac{c^*}{4}\right) + 3p \log \frac{b}{2L\|\psi'\|_{C^0}}.
\end{aligned}$$

To finish, we choose p small enough depending on α, then c^* depending on p. The bound in (9) follows on taking L sufficiently large.

The proof of Proposition 10 is now complete.

4. Proof of Theorem 2

In this section we view our randomly perturbed dynamical system as the Markov chain (X_n, Y_n) defined in Sect. 1, and recall that $P^k((X, Y), \cdot)$ is the k-step transition probability starting from (X, Y). Let $K = K(\delta)$ be as defined at the beginning of Sect. 3, and let $\|\cdot\|_{tv}$ denote the total variation norm of a signed measure.

4.1. Reductions and coupling procedure. Our main proposition is the following.

Proposition 13. *Let $\delta \in (0, 1)$. Then there is a constant $C_0 = C_0(\delta) > 0$ such that the following holds for all L sufficiently large and $\epsilon \in [L^{-1+\delta}, \frac{1}{2}]$: for any (X_0, Y_0), $(X'_0, Y'_0) \in \mathcal{C}$,*

$$\|P^{K+4}((X_0, Y_0), \cdot) - P^{K+4}((X'_0, Y'_0), \cdot)\|_{tv} \leq C_0 L^{-\delta/4}. \tag{18}$$

From (18), it follows easily that for any two Borel probability measures μ_1 and μ_2 on $\mathcal{C}$,

$$\|\mu_1 P^{K+4} - \mu_2 P^{K+4}\|_{tv} \leq \frac{1}{2} C_0 L^{-\delta/4} \|\mu_1 - \mu_2\|_{tv} ,$$

from which the assertion of Theorem 2 follows with $K_1 = K + 4$ and $\sigma = \delta/(8K_1)$.

The rest of Sect. 4.1 contains a proof of Proposition 13 modulo two technical estimates, the proofs of which we postpone to the next subsection.

The reason behind the rapid correlation decay is that the maps F_ω have strong expansion in the x-direction, spreading mass around quickly. To highlight this expansion, we put normalized Lebesgue measure ℓ_0 on $\gamma_0 = [X_0 - \epsilon, X_0 + \epsilon] \times \{Y_0\}$, and for fixed $\omega_2, \omega_3, \ldots$, let $\ell_k = \mathcal{F}^k_* \ell_0$ be the measure obtained by transporting ℓ_0 forward by $\mathcal{F}^k := F_{\omega_k} \circ \cdots \circ F_{\omega_2} \circ F$. Likewise define ℓ'_k using (X'_0, Y'_0) in the place of (X_0, Y_0) for the same sequence $\omega_2, \omega_3, \ldots$. We would like to compare ℓ_k to ℓ'_k. In general, these measures are supported on disjoint sets of curves. The next lemma gives conditions under which they can be coupled.

Lemma 14. *For any* (x, y) *and* $(x, y') \in \mathcal{C}$ *with* $|y - y'| < L^{-1+\frac{1}{2}\delta}$,

$$\|P^2((x, y), \cdot) - P^2((x, y'), \cdot)\|_{tv} < L^{-\frac{1}{2}\delta}.$$

Proof. This follows immediately from (5) together with $\epsilon \geq L^{-1+\delta}$. □

The next three lemmas seek to put as large a fraction of the pushed-forward measures ℓ_k and ℓ'_k as possible into coupling position. The terminology of a horizontal curve crossing a rectangle U of the form $U = I \times [0, b] \subset \mathcal{C}$ is as defined at the beginning of Sect. 3.2. Below we let $\widetilde{F} : [0, 1) \times [0, b] \to \mathbb{R} \times [0, b]$ be the lift of F, i.e., it is defined as in (1) but without the "mod 1" in the first coordinate.

Lemma 15. *Fix* $\eta \in (0, \frac{1}{2}]$, *and let* U *be a connected component of* $\mathcal{C} \backslash (\mathcal{D} \cup B_{1-\eta})$. *Then there is an interval* $J = [a^-, a^+] \subset \mathbb{R}$ *with* $|J| \geq \text{Const.} \cdot L$ *for which the following holds: Let* γ, γ' *be horizontal curves crossing* U, *and let* $h : I \to [0, b]$, $h' : I' \to [0, b]$, *denote the graphing functions of* $\widetilde{F}(\gamma)$, $\widetilde{F}(\gamma')$, *respectively. Then*

$$J = [a^-, a^+] \subset I, I' \subset [a^- - 1, a^+ + 1] ,$$

and on J,

$$\|h - h'\|_{C^0} < L^{-1+\eta}.$$

With obvious modifications, Lemma 15 is valid if $\widetilde{F}$ is replaced by $\widetilde{F}_\omega$ for fixed $\omega \in [-\epsilon, \epsilon]$. Recall that π_1 denotes projection to the first coordinate.

Proof of Lemma 15. First observe that $\widetilde{F}$ maps each of the two horizontal boundaries of U, $I \times \{0\}$ and $I \times \{b\}$, to a horizontal curve of length const $\cdot L$, and it maps each of the two vertical boundaries of U to horizontal segments of length b. Let $[a^-, a^+] = \pi_1 \widetilde{F}(I \times \{0\}) \cap \pi_1 \widetilde{F}(I \times \{b\})$. Since any horizontal curve γ crossing U connects its vertical boundaries, it follows that $\pi_1 \widetilde{F}(\gamma) \supset [a^-, a^+]$ and is contained $[a^- - 1, a^+ + 1]$.

The last assertion in the lemma follows once we check that any vertical line segment connecting $\widetilde{F}(J \times \{0\})$ and $\widetilde{F}(J \times \{b\})$ is expanded by a factor $\geq L^{1-\eta}$ under $\widetilde{F}^{-1}$. This is a simple exercise. □

The next lemma is similar to Proposition 11, with an improved distortion bound.

Lemma 16. *The hypotheses and notation are as in Proposition 11, but we fix $\omega_{K+1}, \omega_{K+2}$ in addition to $\omega_2, \ldots, \omega_K$. Then there exists a constant $\hat{C} = \hat{C}(\delta) > 0$ and a set $\hat{R} \subset [-\epsilon, \epsilon]$ such that*

(a) $\nu^\epsilon([-\epsilon, \epsilon] \backslash \hat{R}) \leq \hat{C} L^{-\delta/4}$, *and*
(b) *on $\hat{R}$ there is a partition $\hat{\mathcal{W}} = \{\hat{W}\}$ into intervals with the following properties:*
 (i) *the set $\{(X_{K+1}, Y_{K+1}) = F_{\omega_{K+1}} \circ \cdots \circ F_{\omega_2} \circ F_\omega(X_0, Y_0) : \omega \in \hat{W}\}$ is a horizontal curve crossing $(-\omega_{K+2}, 1 - \omega_{K+2}) \times [0, b]$; and*
 (ii) *viewing X_{K+1} as a function of $\omega \in \hat{W}$, we have that for $\omega, \omega' \in \hat{W}$,*

$$\left| \frac{\frac{\partial}{\partial \omega} X_{K+1}(\omega)}{\frac{\partial}{\partial \omega} X_{K+1}(\omega')} \right| \leq 1 + \hat{C} L^{-1/2}. \tag{19}$$

Given (X_0, Y_0), we let ℓ_0 be as above, and write $\mathcal{F}^{K+1} = F_{\omega_{K+1}} \circ \cdots \circ F_{\omega_2} \circ F$. Then

$$\hat{\ell}_{K+1} := \mathcal{F}_*^{K+1} \left(\ell_0|_{(X_0 + \hat{R}) \times \{Y_0\}} \right)$$

is supported on a finite number of horizontal curves, each one of which crosses $(-\omega_{K+2}, 1-\omega_{K+2}) \times [0, b]$. Analogous objects using the same sequence $\omega_2, \ldots, \omega_{K+1}$ are defined for (X_0', Y_0'), and are denoted by the same letter with a prime ($'$).

From here on we use C as a generic constant that is allowed to depend only on ψ and on δ; its value may differ from statement to statement but will be increased at most a finite number of times.

Lemma 17. *Given (X_0, Y_0) and (X_0', Y_0'), we fix $\omega_2, \ldots, \omega_{K+2}$, and let U be a component of $\mathcal{C} \backslash ((\mathcal{D} \cup B_{1-\eta}) - (\omega_{K+2}, 0))$ for some $\eta \in (0, 1/2]$. Define*

$$\rho_{K+2,U} = \left(\pi_1 \circ \widetilde{F}_{\omega_{K+2}}\right)_* \left(\hat{\ell}_{K+1}|_U\right),$$

and let $\rho'_{K+2,U}$ be defined analogously. Then

$$\|\rho_{K+2,U} - \rho'_{K+2,U}\|_{tv} \leq C L^{-\min\left\{\frac{1}{2}, 1-2\eta, \frac{\delta}{4}\right\}}. \tag{20}$$

We first complete the proof of Proposition 13, postponing the proofs of Lemmas 16 and 17 to the next subsection so as not to disrupt the flow of ideas.

Proof of Proposition 13 assuming Lemmas 16 and 17. Fixing arbitrary $\omega_2, \ldots, \omega_{K+2} \in [-\epsilon, \epsilon]$, we compare the conditional probabilities

$$P^{K+4}(((X_0, Y_0), \cdot)|\omega_2, \ldots, \omega_{K+2}) \quad \text{and}$$
$$P^{K+4}\left(\left(\left(X_0', Y_0'\right), \cdot\right)|\omega_2, \ldots, \omega_{K+2}\right)$$

as follows:

Let ℓ_k, $k \leq K + 2$, be as above; the measures ℓ_k' are defined similarly for (X_0', Y_0').

Lemma 16 says that omitting a fraction $\leq C L^{-\delta/4}$ of ℓ_{K+1} and ℓ'_{K+1}, we obtain $\hat{\ell}_{K+1}$ and $\hat{\ell}'_{K+1}$, each supported on a collection of horizontal curves that cross $(-\omega_{K+2}, 1 - \omega_{K+2}) \times [0, b]$, with nearly constant densities (in the sense of (19)) on each curve. We retain only those parts of $\hat{\ell}_{K+1}$ and $\hat{\ell}'_{K+1}$ that lie outside of $B_{1-\eta} - (\omega_{K+2}, 0)$ (for some η to be determined), thereby discarding another fraction of total mass $\leq CL^{-\eta}$.

Let U be a component of $\mathcal{C} \backslash ((\mathcal{D} \cup B_{1-\eta}) - (\omega_{K+2}, 0))$, and let $\hat{\ell}_{K+1,U}$ and $\hat{\ell}'_{K+1,U}$ be the restrictions of $\hat{\ell}_{K+1}$ and $\hat{\ell}'_{K+1}$ to U. Applying Lemmas 15 and 17, we deduce that there are parts $\check{\ell}_{K+2,U}$ and $\check{\ell}'_{K+2,U}$ of $(\widetilde{F}_{\omega_{K+2}})_*(\hat{\ell}_{K+1,U})$ and $(\widetilde{F}_{\omega_{K+2}})_*(\hat{\ell}'_{K+1,U})$ respectively with the following properties:

(i) $(\pi_1)_*\check{\ell}_{K+2,U} = (\pi_1)_*\check{\ell}'_{K+2,U}$;
(ii) restricted to any vertical line, the support of $\check{\ell}_{K+2,U} \cup \check{\ell}'_{K+2,U}$ is contained in an interval of length $< L^{-1+\eta}$; and
(iii) the total mass of $(\widetilde{F}_{\omega_{K+2}})_*(\hat{\ell}_{K+1,U}) - \check{\ell}_{K+2,U}$ is $\leq CL^{-\min\{\frac{1}{2}, 1-2\eta, \delta/4\}}$, with the analogous estimate for $\check{\ell}'_{K+2,U}$.

Choosing $\eta < \frac{\delta}{2}$, we have, by Lemma 14, that the measures $\check{\ell}_{K+2}$ and $\check{\ell}'_{K+2}$ are in coupling position. That is to say, in the next 2 steps, a fraction of $\geq 1 - L^{-\frac{1}{2}\delta}$ of them will coincide.

Finally, the total mass that remains uncoupled after $K + 4$ steps is

$$\leq CL^{-\min\{\delta/4, \eta, 1-2\eta, 1/2\}}.$$

Assuming $\eta \leq \frac{1}{3}$ and $\delta < 1$, this is $< CL^{-\delta/4}$, concluding the proof of Proposition 13. $\square$

4.2. Technical estimates. Lemma 16 is an extension of Proposition 11, and we follow the notation there.

Proof of Lemma 16. We define $\hat{W}$ and $\hat{R}$ on one $W \in \mathcal{W}$ at a time, where $\mathcal{W}$ is as in Proposition 11. Let γ be the image of W under $\omega \mapsto (X_K(\omega), Y_K(\omega))$. Mirroring the proof of Proposition 11, we let $\hat{\mathcal{R}}_\gamma$ be the collection of subsegments $\hat{\gamma}$ of $\gamma \backslash (B_{3/4} \cup \mathcal{D})$ with the property that $F_{\omega_{K+1}}(\hat{\gamma})$ crosses $(-\omega_{K+2}, 1 - \omega_{K+2}) \times [0, b]$. The pullbacks of elements of $\hat{\mathcal{R}}_\gamma$ to W are elements of $\hat{\mathcal{W}}$ and their union is $W \cap \hat{R}$.

The additional deletion in this last step can be estimated as follows: Let W and γ be as in the last paragraph. From γ, we culled $\#C'_\psi$ subsegments due to $B_{\frac{3}{4}}$, and at most $2(\#C'_\psi + 1)$ subsegments due to the failure of their images to cross $(-\omega_{K+2}, 1 - \omega_{K+2}) \times [0, b]$. This leads to a deletion from γ of total length

$$\leq \#C'_\psi \cdot 2D_1 L^{-\frac{1}{4}} + 2\left(\#C'_\psi + 1\right) \cdot L^{-\frac{3}{4}}$$

which when pulled back to W amounts to a fraction $\leq$ Const. $\cdot\, L^{-\frac{1}{4}}$ of W.

Next we estimate the distortion of X_{K+1} on $\hat{W}$, where $\hat{W} \subset W$ are as above. In the notation of the proof of Proposition 11, we will check the distortion of

$$\check{F}_i : \hat{I}_{i-1} \to \hat{I}_i\,, \quad \text{where} \quad \hat{I}_i = \pi_1 \mathcal{F}^i \left(\left(X_0 + \hat{W}\right) \times \{Y_0\}\right),$$

for $1 \leq i \leq K + 1$.

Consider first $1 \le i \le K$. Let $\{I_i\}_{1\le i\le K}$ be as in Proposition 11. Then $\check{F}_K \circ \cdots \check{F}_{i+1}$ maps I_i onto I_K and $\hat{I}_i$ onto $\hat{I}_K$; moreover, its distortion is bounded by a constant independent of L (by Proposition 11). Since $\hat{I}_K \cap B_{3/4} = \emptyset$ and $|\hat{I}_{K+1}| = 1$ by construction, we have $|\hat{I}_K| \le L^{-3/4}$, and the distortion bound above gives

$$\frac{|\hat{I}_i|}{|I_i|} \le \text{Const.}\ \frac{|\hat{I}_K|}{|I_K|} \le \text{Const.}\ L^{-3/4}.$$

Applying this to either of the estimates (12) or (13) (depending on the length of γ_{i-1} and $F_{\omega_i}(\gamma_{i-1})$; see the proof of Proposition 11), it immediately follows that

$$\left|\log \frac{(\check{F}_i)'(x_1)}{(\check{F}_i)'(x_2)}\right| \le \text{Const.}\ L^{-3/4}$$

for any $x_1, x_2 \in \hat{I}_{i-1}$.

For the case $i = K+1$, applying our estimate for $|\hat{I}_K|$ yields

$$\left|\log \frac{(\check{F}_{K+1})'(x_1)}{(\check{F}_{K+1})'(x_2)}\right| \le \frac{\sup_{z\in\hat{I}_K} |\check{F}''_{K+1}(z)|}{\inf_{z\in\hat{I}_K} |\check{F}'_{K+1}(z)|} \cdot |\hat{I}_K| \le \frac{2\hat{D}L}{(L^{\frac{3}{4}})^2} \le \text{Const.}\ L^{-1/2}$$

for any $x_1, x_2 \in \hat{I}_K$. This completes the distortion bound. □

The following comments may help motivate our proof of Lemma 17 below. The measure $(\widetilde{F}_{\omega_{K+2}})_*(\hat{\ell}_{K+1}|_U)$ is supported on a collection Γ_U of horizontal curves of length $O(L)$, and $(\widetilde{F}_{\omega_{K+2}})_*(\hat{\ell}'_{K+1}|_U)$ is supported on a collection Γ'_U. As these curves are quite close in vertical distance, one might consider coupling the measure on each $\check{\gamma} \in \Gamma_U$ to that on some $\check{\gamma}' \in \Gamma'_U$ as in Lemma 14. This does not work well, because the weights carried by individual curves vary. That is why we compare only the π_1-projections of the measures $(\widetilde{F}_{\omega_{K+2}})_*(\hat{\ell}_{K+1}|_U)$ and $(\widetilde{F}_{\omega_{K+2}})_*(\hat{\ell}'_{K+1}|_U)$. A second point is that the pushforward density on each $\check{\gamma}$ (or $\check{\gamma}'$) is nowhere close to being constant along the entire length of the curve, as $\partial_x \widetilde{F}_{\omega_{K+2}}$ varies considerably over the length of its pre-image.

Proof of Lemma 17. Let U be fixed, and let $J = [a^-, a^+]$ be given by Lemma 15. Let γ be a horizontal curve of the form $\omega \mapsto (X_{K+1}(\omega), Y_{K+1}(\omega)), \omega \in \hat{W}$ for some $\hat{W} \in \hat{\mathcal{W}}$, where $\hat{\mathcal{W}}$ is as in Lemma 16. We fix $\bar{x} \in J$, and estimate the density of $(\pi_1 \circ \widetilde{F}_{\omega_{K+2}})_*(\hat{\ell}_{K+1}|_{\gamma\cap U})$ at $\bar{x}$ as follows. Let $(x_1, y_1) \in \gamma \cap U$ be such that $\pi_1 \widetilde{F}_{\omega_{K+2}}(x_1, y_1) = \bar{x}$. Using the line $\{y = b\}$ as a reference, we let $(\mathbf{x}, b) \in U$ be such that $\pi_1 \widetilde{F}_{\omega_{K+2}}(\mathbf{x}, b) = \bar{x}$. Let ∂_1 and ∂_2 denote unit vectors in the x- and y-directions, and suppose $\partial_1 + s\partial_2$ is tangent to γ at (x_1, y_1). We claim that

$$\left|\frac{\pi_1(d\widetilde{F}_{\omega_{K+2}})_{(x_1,y_1)}(\partial_1 + s\partial_2)}{\pi_1(d\widetilde{F}_{\omega_{K+2}})_{(\mathbf{x},b)}\partial_1} - 1\right| < CL^{-1+2\eta} \tag{21}$$

for some C independent of γ.

The assertion in Lemma 17 can be deduced from (21) as follows. First,

$$\pi_1(d\widetilde{F}_{\omega_{K+2}})_{(\mathbf{x},b)}\partial_1 = f'(\mathbf{x} + \omega_{K+2}),$$

and the density of $(\pi_1)_*(\hat{\ell}_{K+1}|_{\gamma\cap U})$ is $c_\gamma(1 \pm CL^{-\frac{1}{2}})$ on $\pi_1(U)$ where $c_\gamma = \hat{\ell}_{K+1}(\gamma)$ (Lemma 16). These together with (21) imply that the density of $(\pi_1 \circ \widetilde{F}_{\omega_{K+2}})_*(\hat{\ell}_{K+1}|_{\gamma\cap U})$ at $\bar{x}$ is, up to a multiplicative factor of $(1 \pm CL^{-\min\{\frac{1}{2},1-2\eta\}})$, equal to $c_\gamma/|f'(\mathbf{x}+\omega_{K+2})|$. Summing over all γ from $\hat{W} \in \hat{\mathcal{W}}$ and recalling that $\sum c_\gamma = \hat{\ell}_{K+1}(\mathcal{C}) > 1 - CL^{-\delta/4}$ (Lemma 16), it follows that up to a multiplicative factor of $(1 \pm CL^{-\min\{\frac{1}{2},1-2\eta,\frac{\delta}{4}\}})$, the density of $(\pi_1 \circ \widetilde{F}_{\omega_{K+2}})_*(\hat{\ell}_{K+1}|_U)$ at $\bar{x}$ is $1/|f'(\mathbf{x}+\omega_{K+2})|$. An analogous estimate holds for $(\pi_1 \circ \widetilde{F}_{\omega_{K+2}})_*(\hat{\ell}'_{K+1}|_U)$. Thus the density of the signed measure $\rho_{K+2,U} - \rho'_{K+2,U}$ at $\bar{x}$ has absolute value $\leq CL^{-\min\{\frac{1}{2},1-2\eta,\frac{\delta}{4}\}}/|f'(\mathbf{x}+\omega_{K+2})|$. Integrating over $\bar{x} \in J$ and pulling the integral back to $(f^{-1}(J+b) \cap \pi_1(U)) - \omega_{K+2}$ along the map $\mathbf{x} \mapsto \bar{x} = f(\mathbf{x}+\omega_{K+2}) - b$, the desired conclusion follows.

It remains to prove (21). From the definition of F, and the fact that $(\mathbf{x}+\omega_{K+2}, b) \notin B_{1-\eta}$, the left side of (21) is

$$\leq L^{-1+\eta} \cdot (|f'(x_1+\omega_{K+2}) - f'(\mathbf{x}+\omega_{K+2})| + |s|) \tag{22}$$

It is easy to check that $(\widetilde{F}_{\omega_{K+2}}|_U)^{-1}(\{\bar{x}\} \times [0,b])$ is a *vertical* curve, i.e., it is the graph of a function g from $[0,b]$ in the y-axis to the x-axis with $|g'| \leq L^{-1+\eta}$. This implies $|x_1 - \mathbf{x}| \leq L^{-1+\eta}$, so

$$|f'(x_1+\omega_{K+2}) - f'(\mathbf{x}+\omega_{K+2})| \leq \|\psi''\|_{C^0} L \cdot L^{-1+\eta} = \|\psi''\|_{C^0} L^\eta.$$

Finally, $|s| < 1$ since γ is a horizontal curve, and (21) is proved. □

5. Proof of Theorems 3 and 4

The proofs of Theorems 3 and 4 shadow those of Theorems 1 and 2 closely with a few differences. In both proofs, the main work is in showing that the additional assumption (H3) permits us to "grow" a horizontal curve of length $\sim L^{-2+\delta}$ to one of length $\sim L^{-1+\delta}$. This is discussed in Sect. 5.1. Once this is achieved, earlier techniques can be brought to bear. Theorem 3 is proved in Sect. 5.2, and Theorem 4 is proved in Sect. 5.3.

Throughout this section, α and δ are fixed, and we assume (H1), (H2) and $(H3)_{c_0}$ for some fixed $c_0 > 0$ (see Sect. 1). We assume further that L is sufficiently large, and consider $\epsilon \in [L^{-2+\delta}, L^{-3/4}]$, appealing to Theorems 1 and 2 for $\epsilon > L^{-3/4}$. Let K be as defined at the beginning of Sect. 3.

5.1. Growing horizontal curves of length $L^{-2+\delta}$. We have shown in Sect. 3 that any horizontal curve of length $\geq L^{-1+\delta}$ will, in K steps, wrap around $\mathcal{C}$ many times under $F^K_{\underline{\omega}}$ for any $\underline{\omega}$. Here, since randomization produces only segments of length $2L^{-2+\delta}$, we seek to grow these short segments to length $\geq L^{-1+\delta}$. We will show that this can be done in a few steps, provided we randomize "at the right times".

First we state a lemma that is a direct consequence of (H3). Let $c = c_\psi \ll c_0$ where c_0 is as in (H3); how small c has to be will become clear as we go along. Following the notation just before the statement of Theorem 3, we consider the following subdivision of $\mathbb{S}^1$:

$$\mathcal{B} = \mathcal{N}_{\sqrt{\frac{c}{L}}}\left(C'_\psi\right), \quad \mathcal{I} = \mathcal{N}_c\left(C'_\psi\right) \backslash \mathcal{B}, \quad \text{and} \quad \mathcal{G} = \mathbb{S}^1 \backslash (\mathcal{B} \cup \mathcal{I}).$$

Lemma 18. *Let (X_0, Y_0) be such that $X_0 \in \mathcal{B} \cup \mathcal{I}$, and assume $X_1 \in \mathcal{B}$. Then $X_2 \in \mathcal{G}$.*

Proof. Let $X_1 = f(X_0 + \omega_1) - Y_0 \pmod 1$, $X_2 = f(X_1 + \omega_2) - Y_1 \pmod 1$, and let $\hat{x}_0, \hat{x}_1 \in C'_\psi$ (possibly $\hat{x}_0 = \hat{x}_1$) be such that $|X_0 - \hat{x}_0| < c$ and $|X_1 - \hat{x}_1| < \sqrt{\frac{c}{L}}$. Observe that for L large, we have

$$|f(X_1 + \omega_2) - f(\hat{x}_1)| < \frac{1}{2} L \|\psi''\| \left(\sqrt{\frac{c}{L}} + L^{-3/4} \right)^2 < \|\psi''\| c \ ,$$

and $|Y_1 - b\hat{x}_0| = |b(X_0 + \omega_1) - b\hat{x}_0| < 2c$. Thus

$$|(f(\hat{x}_1) - b\hat{x}_0) - (f(X_1 + \omega_2) - Y_1)| < (2 + \|\psi''\|)c.$$

Since by (H3), $f(\hat{x}_1) - b\hat{x}_0 \notin \mathcal{N}_{c_0}(C'_\psi)$, it follows that $X_2 \notin \mathcal{N}_{\frac{1}{2}c_0}(C'_\psi)$. □

The following notation will be useful: Given any function $\tau : \mathcal{C} \times [-\epsilon, \epsilon]^{\mathbb{N}} \to \mathbb{Z}^+$, we define for each $(X_0, Y_0) \in \mathcal{C}$ and $\underline{\omega} \in [-\epsilon, \epsilon]^{\mathbb{N}}$ a sequence of sets $\zeta_k = \zeta_k((X_0, Y_0), \underline{\omega})$ for $k = 1, 2, \ldots$ as follows:

For $k < \tau$, $\zeta_k = \{F^k_{\underline{\omega}}((X_0, Y_0))\} = \{(X_k, Y_k)\}$.
For $k = \tau$, $\zeta_k = F([X_{k-1} - \epsilon, X_{k-1} + \epsilon] \times \{Y_{k-1}\})$.
For $k > \tau$, $\zeta_k = F_{\omega_k}(\zeta_{k-1})$.

That is to say, for $k < \tau$, ζ_k is the image of (X_0, Y_0) under $F^k_{\underline{\omega}}$. At time τ, we randomize, and for $k \geq \tau$, ζ_k is a curve, or a disjoint union of continuous curves, which can be thought of as parametrized by $\omega \in [-\epsilon, \epsilon]$. Abusing notation slightly, we will write this parametrization as $\omega \mapsto \zeta_k(\omega)$.

Let $m \leq 1$ be such that

$$\begin{aligned} &\text{for all } x \in [0,1) \backslash \mathcal{N}_{\frac{9}{10}c}\left(C'_\psi\right), && |f'(x)| > mL \\ \text{and} \quad &\text{for all } x \in [0,1) \backslash \mathcal{N}_{\frac{9}{10}\sqrt{\frac{c}{L}}}\left(C'_\psi\right), && |f'(x)| > mL^{-\frac{1}{2}}. \end{aligned}$$

Since $\frac{1}{10}\sqrt{\frac{c}{L}} \gg L^{-3/4} \geq \epsilon$, we may view $[x - \epsilon, x + \epsilon]$ as essentially contained in $\mathcal{I}$ for $x \in \mathcal{I}$, and similarly for $x \in \mathcal{G}$. Recall that F is discontinuous at $\mathcal{D} = \{x = 0\}$. We may assume $\mathcal{I}$ is away from 0.

Lemma 19. *There exist stopping times $\tau_0 \leq \tau_1$ with $1 \leq \tau_0, \tau_1 \leq 4$ such that for each (X_0, Y_0) and $\underline{\omega}$, if ζ_k is constructed as above using $\tau = \tau_0$, then*

(i) *ζ_{τ_1} is the union of at most two horizontal curves, and*
(ii) *its total length is $\geq 2m^2 L^{-1+\delta}$.*

Proof. Let (X_0, Y_0) and $\underline{\omega}$ be given.

Case 1. $X_0 \in \mathcal{G}$. Letting $\tau_0 = \tau_1 = 1$, we have $|\zeta_1| > mL(2\epsilon) \geq 2mL^{-1+\delta}$; ζ_1 may have one or two components depending on whether $(X_0 - \epsilon, X_0 + \epsilon) \cap \{0\} = \emptyset$.

Case 2. $X_0 \in \mathcal{I}$. Here if $\widetilde{\zeta}_1 = F([X_0 - \epsilon, X_0 + \epsilon] \times \{Y_0\})$, then we are guaranteed only $|\widetilde{\zeta}_1| > 2mL^{-\frac{3}{2}+\delta}$. We divide into the following subcases depending on properties of $\widetilde{\zeta}_1$.

Case 2a. $|\widetilde{\zeta}_1| \geq L^{-1+\delta}$. In this case, we may take $\tau_0 = \tau_1 = 1$.

Case 2b. $|\widetilde{\zeta}_1|<L^{-1+\delta}$ and $\widetilde{\zeta}_1 \cap (\mathcal{I} \cup \mathcal{G}) \neq \emptyset$. Here $\tau_0=1$ and $\tau_1=2$ gives the result that $\zeta_2 = F_{\omega_2}(\widetilde{\zeta}_1)$ is a horizontal curve of length $> 2m^2 L^{-1+\delta}$.

Case 2c. $|\widetilde{\zeta}_1|<L^{-1+\delta}$ and $\widetilde{\zeta}_1 \subset \mathcal{B}$. Then $X_1 \in \mathcal{B}$, and by Lemma 18, $X_2 \in \mathcal{G}$. We are thus in the situation of Case 1, and $\tau_0 = \tau_1 = 3$ gives the desired result.

Case 3. $X_0 \in \mathcal{B}$. Here if $X_1 \in \mathcal{G}$ or in $\mathcal{I}$, we do as in Cases 1 or 2, with (X_1, Y_1) playing the role of (X_0, Y_0), resulting in $2 \le \tau_0, \tau_1 \le 4$. If $X_1 \in \mathcal{B}$, then by Lemma 18 we again have $X_2 \in \mathcal{G}$ and $\tau_0 = \tau_1 = 3$ works. □

Observe that τ_0 and τ_1 are *bona fide* stopping times: For (X_0, Y_0) with $X_0 \in \mathcal{G} \cup \mathcal{I}$, τ_0 and τ_1 depend only on (X_0, Y_0); while for $X_0 \in \mathcal{B}$, $\tau_0, \tau_1 \ge 2$ and depend on (X_0, Y_0) and ω_1. Observe also that all cases reduce to Cases 1, 2a or 2b after suitable time delays, and $X_{\tau_0} \in (\mathcal{I} \cup \mathcal{G})$ in all cases.

Next we prove the analog of Proposition 11.

Proposition 20. *Given any $\delta > 0$, there exist $L^* = L^*(\delta) \ge 1$ and $C = C(\delta) \ge 1$ such that the following hold for all $L \ge L^*$ and $\epsilon \in [L^{-2+\delta}, L^{-3/4}]$. Fix $(X_0, Y_0) \in \mathcal{C}$ and $\omega_i \in [-\epsilon, \epsilon]$, $i = 1, 2, \dots, K+4$. Let τ_0 be the stopping time given by Proposition 19. Randomizing at step τ_0, we let $\omega \mapsto \zeta_k(\omega) = (X_k(\omega), Y_k(\omega))$ be the parametrization defined earlier for $k \ge \tau_0$. Then there exists a set $R \subset [-\epsilon, \epsilon]$ such that*

(a) $\nu^\epsilon([-\epsilon, \epsilon]\backslash R) \le C L^{-\delta/4}$, *and*
(b) *on R there is a partition $\mathcal{W} = \{W\}$ into intervals with the following properties:*
  (i) *the set $\{\zeta_{K+4}(\omega), \omega \in W\}$ is a horizontal curve that crosses $\mathcal{C}$; and*
  (ii) *for $\omega, \omega' \in W$,*

$$\left| \frac{\frac{\partial}{\partial\omega} X_{K+4}(\omega)}{\frac{\partial}{\partial\omega} X_{K+4}(\omega')} \right| \le C. \tag{23}$$

Proof. Let (X_0, Y_0) and $\underline{\omega}$ be given. We follow the steps in the proof of Proposition 11.
(A) We start the construction in Proposition 11 with elements of Γ_{τ_1} playing the role of γ_0, where Γ_{τ_1} is defined as follows: If ζ_{τ_0} is connected, which is true except possibly in Case 1 of Lemma 19, we let $\Gamma_{\tau_0} = \{\zeta_{\tau_0}\}$. In Case 1, we keep only those components of $[X_{\tau_0-1} - \epsilon, X_{\tau_0-1} + \epsilon] \times \{Y_{\tau_0-1}\}\backslash\mathcal{D}$ that have length $> \frac{1}{m} L^{-\delta/4} \cdot 2\epsilon$, and put their F-images into Γ_{τ_0}. This completes the definition of $\Gamma_{\tau_1} = \Gamma_{\tau_0}$ in Cases 1 and 2a. In Case 2b, we let $\Gamma_{\tau_1} = \{\zeta_{\tau_1}\}$. Observe that Γ_{τ_1} so constructed consists of one or two horizontal curves, the length of each being $\ge \min(2L^{-1+\frac{3\delta}{4}}, 2m^2 L^{-1+\delta})$. Assuming L is large enough to absorb the factor m^2, one checks that the construction of $\Gamma_{\tau_1+1}, \Gamma_{\tau_1+2}, \dots, \Gamma_{K+4}$ can proceed exactly as before, with all the curves in Γ_{K+4} horizontal and crossing $(0, 1) \times [0, b]$.
(B) One needs only to be concerned with distortion in the first steps. For illustration, consider Case 2b in Lemma 19. Here $\tau_1 = \tau_0 + 1$. We let

$$I_{\tau_0-1} = \left[X_{\tau_0-1} - \epsilon, X_{\tau_0-1} + \epsilon\right], \quad I_{\tau_0} = \pi_1(\zeta_{\tau_0}) \quad \text{and} \quad I_{\tau_1} = \pi_1(\zeta_{\tau_1}).$$

Observe that $I_{\tau_0-1}, I_{\tau_0} \subset (\mathcal{I} \cup \mathcal{G})$, and $|I_{\tau_0-1}|, |I_{\tau_0}| < L^{-1+\delta}$. This implies that for $i = \tau_0, \tau_1$ and $x_1, x_2 \in I_{i-1}$, (12) becomes

$$\left| \log \frac{(\check{F}_i)'(x_1)}{(\check{F}_i)'(x_2)} \right| \le \frac{\sup_{z \in I_{i-1}} |\check{F}_i''(z)|}{\inf_{z \in I_{i-1}} |\check{F}_i'(z)|} \cdot |I_{i-1}| \le \frac{2\hat{D}L}{\frac{1}{2} m L^{\frac{1}{2}}} \cdot 2L^{-1+\delta} = o(1). \tag{24}$$

The main change in (C) is the deletion at step τ_0 in Case 1. The fraction deleted is $\le \frac{1}{m} L^{-\delta/4}$, which is allowed. □

5.2. Estimation of Lyapunov exponents.

Proposition 21. *Assume the hypothesis of Theorem 3, and let* $\bar{K} = K + 4$*. Then given* $\delta > 0$*, there exists* L^* *such that for all* $L > L^*$ *and* $\epsilon \in [L^{-2+\delta}, L^{-3/4}]$*, the estimate*

$$\int \log \|(dF_{\omega_{\bar{K}+3}})_{(X_{\bar{K}+2},Y_{\bar{K}+2})} u_{\theta_{\bar{K}+2}}\| \, (d\nu^{\epsilon})^{\otimes 5}(\omega_1, \omega_2, \omega_3, \omega_4, \omega_{\bar{K}+1}) \geq \alpha \log L \tag{25}$$

holds for any $(X_0, Y_0) \in \mathcal{C}$*,* $\theta_0 \in [0, 2\pi)$*, and any given* $\omega_5, \ldots, \omega_{\bar{K}}, \omega_{\bar{K}+2}, \omega_{\bar{K}+3} \in [-\epsilon, \epsilon]$.

The proof follows closely that of Proposition 10. We focus on the modifications.

Proof. Given (X_0, Y_0) and $\omega_i \in [-\epsilon, \epsilon]$, $5 \leq i \leq \bar{K}$, we let $\Delta_4 = \cup_{\ell=1}^4 \Delta_{4,\ell}$ where

$$\Delta_{4,\ell} = \{(\omega_1, \ldots, \omega_4) : \tau_0((X_0, Y_0), \omega_1, \ldots, \omega_{\ell-1}) = \ell \text{ and } \omega_\ell \in R\}$$

where R is the set of "good ω" given by part (a) in Proposition 20. It follows from Proposition 20 that $(\nu^\epsilon)^{\otimes 4}(\Delta_4) > 1 - p$ for L large enough. For each $\omega_{\bar{K}+1}$, we define

$$G_1(\omega_{\bar{K}+1}) = \Big\{(\omega_1, \ldots, \omega_4) \in \Delta_4 : X_{\bar{K}} + \omega_{\bar{K}+1}, X_{\bar{K}+1} + \omega_{\bar{K}+2}, X_{\bar{K}+2} + \omega_{\bar{K}+3} \notin B^1_{c^*} \quad \text{and} \quad X_{\bar{K}} + \omega_{\bar{K}+1} \notin B^2_{c^*}\Big\}.$$

The same argument as before gives $(\nu^\epsilon)^{\otimes 4}(G_1(\omega_{\bar{K}+1})) > 1 - 2p$.

The sets G_1 and G_2 are defined analogously as in Proposition 10, and the measure of $G_2^c(\omega_1, \ldots, \omega_4)$ is estimated similarly. The only difference occurs in the analog of (15), where assuming only $\epsilon \geq L^{-2+\delta}$, we can guarantee only

$$\nu^\epsilon \left(G_2^c(\omega_1, \ldots, \omega_4)\right) < \frac{b}{c^*} L^{\sigma-\delta} ,$$

requiring that we take $\sigma < \delta$.

With these modifications, the proof in Sect. 3.3 goes through. □

5.3. Proof of Theorem 4. Let $\bar{K} = K + 4$, where K is as in the beginning of Sect. 3. It suffices to show the following.

Proposition 22. *Let* $c_0 > 0$ *be fixed. Then given* $\delta \in (0, 1)$*, there is a constant* $C = C(\delta) > 0$ *such that the following holds for all* L, ϵ *and a satisfying the conditions in Theorem 4: For any* (X_0, Y_0)*,* $(X_0', Y_0') \in \mathcal{C}$*,*

$$\| P^{\bar{K}+5}\left((X_0, Y_0), \cdot\right) - P^{\bar{K}+5}\left(\left(X_0', Y_0'\right), \cdot\right) \|_{tv} \leq C L^{-\delta/4}.$$

The proof of this proposition parallels that of Proposition 13. As in the previous subsection, we focus on the more substantive differences, of which there are two.

(1) Adapting to variable randomization times

Instead of conditioning on $\omega_2, \ldots, \omega_{K+2}$ as was done in Proposition 13, here we compare the conditional probabilities

$$P^{\bar{K}+5}\left(((X_0, Y_0), \cdot) \,|g_{(X_0,Y_0)}\left(\bar{\omega}_1, \bar{\omega}_2, \bar{\omega}_3\right), \omega_5, \ldots, \omega_{\bar{K}+3}\right)$$

and $$P^{\bar{K}+5}\left(\left(\left(X_0', Y_0'\right), \cdot\right) |g_{(X_0',Y_0')}\left(\bar{\omega}_1, \bar{\omega}_2, \bar{\omega}_3\right), \omega_5, \ldots, \omega_{\bar{K}+3}\right)$$

for each sequence $(\bar\omega_1, \bar\omega_2, \bar\omega_3; \omega_5, \ldots, \omega_{\bar K+3})$ where $g_{(X_0,Y_0)}(\cdots)$ is defined as follows: From Sect. 5.1, there exists $\ell \le 4$ such that

$$\tau_0((X_0, Y_0), \bar\omega_1, \ldots, \bar\omega_{\ell-1}) = \ell.$$

We define $g_{(X_0,Y_0)}(\bar\omega_1, \bar\omega_2, \bar\omega_3)$ to be the triple $(\omega_i)_{1\le i\le 4, i\ne\ell}$ where

$$\omega_i = \bar\omega_i \text{ for } i < \ell \quad \text{and} \quad \omega_i = \bar\omega_{i-1} \text{ for } i > \ell\ ;$$

that is, we condition on ω_i, $i = 1, \ldots, 4$, $i \ne \ell$, in addition to $\omega_5, \ldots, \omega_{\bar K+3}$. Quantities for (X_0', Y_0') are defined similarly.

Definitions of the measures ℓ_k, ℓ_k' are modified accordingly: We let ℓ_{τ_0-1} be the uniform distribution on $[X_{\tau_0-1} - \epsilon, X_{\tau_0-1} + \epsilon] \times \{Y_{\tau_0-1}\}$, define $\ell_{\tau_0} = F_*\ell_{\tau_0-1}$, and for $k > \tau_0$, let

$$\ell_k = (F_{\omega_k} \circ \cdots \circ F_{\omega_{\tau_0+1}})_*\ell_{\tau_0}.$$

An argument similar to that of Lemma 16 extends Proposition 20 to give a “good set” $\hat R$ with the property that if

$$\hat\ell_{\bar K+1} := \left(F_{\omega_{\bar K+1}} \circ \cdots \circ F_{\omega_{\tau_0+1}} \circ F\right)_* \left(\ell_{\tau_0-1}|_{\left(\{X_{\tau_0-1}\}+\hat R\right)\times\{Y_{\tau_0-1}\}}\right),$$

then

(i) $\hat\ell_{\bar K+1}(\mathcal C) > 1 - CL^{-\delta/4}$,
(ii) $\hat\ell_{\bar K+1}$ is supported on a finite union of horizontal curves γ, each one of which crosses $(-\bar\omega_{\bar K+2}, 1 - \bar\omega_{\bar K+2})$, and
(iii) for each such γ, the density of $(\pi_1)_*(\hat\ell_{\bar K+1}|_\gamma)$ is $c_\gamma(1 \pm CL^{-1/2})$, where $c_\gamma = \hat\ell_{\bar K+1}(\gamma)$.

The same construction applies to (X_0', Y_0'), yielding $\hat\ell'_{\bar K+1}$ with the same properties.

(2) Adapting to smaller ϵ

Next we fix a component U of $\mathcal C\backslash((\mathcal D \cup B_{1-\eta}) - (\omega_{\bar K+2}, 0))$, and seek to compare $(F_{\omega_{\bar K+2}})_*(\hat\ell_{\bar K+1}|_U)$ and $(F_{\omega_{\bar K+2}})_*(\hat\ell'_{\bar K+1}|_U)$. Here we encounter the second major difference: The analog of Lemma 14 in the setting of Theorem 4 is

Lemma 23. *For any* $(x, y), (x, y') \in \mathcal C$ *with* $|y - y'| < L^{-2+\delta/2}$, *we have*

$$\|P^2((x, y), \cdot) - P^2((x, y'), \cdot)\|_{tv} \le L^{-\delta/2}.$$

The supports of $(F_{\omega_{\bar K+2}})_*(\hat\ell_{\bar K+1}|_U)$ and $(F_{\omega_{\bar K+2}})_*(\hat\ell'_{\bar K+1}|_U)$ are guaranteed, by Lemma 15, to be $< L^{-1+\delta/2}$ in vertical distance, not enough to apply Lemma 23. To bring these measures into coupling position, we iterate one more time, using $F_{\omega_{\bar K+3}}$ (the last ω in the sequence on which we condition at the beginning of the proof). To fix some notation, we lift $B_{1-\eta}$ and $\mathcal D$ to $\mathbb R \times [0, b]$, letting $\widetilde B_{1-\eta} = B_{1-\eta} + (\mathbb Z, 0)$ and $\widetilde{\mathcal D} = \mathcal D + (\mathbb Z, 0)$. We also extend the domain of definition of $\widetilde F$ to $\mathbb R \times [0, b]$, by letting $\widetilde F(x, y) = \widetilde F(x+1, y)$. The following is the replacement of Lemma 15.

Lemma 24. *Let U and J be as in Lemma 15, and let $\widetilde{U}$ be a component of $(\mathbb{R} \times [0,b])\backslash(\widetilde{B}_{1-\eta} \cup \widetilde{\mathcal{D}})$ contained in $J \times [0,b]$. Then there exists $\widetilde{J} = [\widetilde{a}^-, \widetilde{a}^+] \subset \mathbb{R}$ with $|\widetilde{J}| \geq const \cdot L$ for which the following hold: Let γ, γ' be horizontal curves crossing U, so that $\widetilde{F}(\widetilde{F}(\gamma|_U)|_{\widetilde{U}})$ and $\widetilde{F}(\widetilde{F}(\gamma'|_U)|_{\widetilde{U}})$ are horizontal curves with graphing functions $\widetilde{h} : \widetilde{I} \to [0,b]$ and $\widetilde{h}' : \widetilde{I}' \to [0,b]$ respectively. Then*

$$\widetilde{J} = [\widetilde{a}^-, \widetilde{a}^+] \subset \widetilde{I}, \widetilde{I}' \subset [\widetilde{a}^- - 1, \widetilde{a}^+ + 1],$$

and on $\widetilde{J}$,

$$\|\widetilde{h} - \widetilde{h}'\|_{C^0} < 2L^{-2+2\eta}.$$

Only the last assertion is new, and it follows from a simple computation. To complete the proof of Proposition 22, we need the following replacement of Lemma 17:

Lemma 25. *Let $\omega_{\bar{K}+2}, \omega_{\bar{K}+3}$ be given, and let $\ell_{\bar{K}+1}$ and $\ell'_{\bar{K}+1}$ be as above. For $\eta \in (0, \frac{1}{2}]$, let U be a component of $\mathcal{C}\backslash((\mathcal{D} \cup B_{1-\eta}) - (\omega_{\bar{K}+2}, 0))$, J as in Lemma 15, and let $\widetilde{U}$ be a component of $(\widetilde{\mathcal{D}} \cup \widetilde{B}_{1-\eta}) - (\omega_{\bar{K}+3}, 0))$ contained in $J \times [0,b]$. Define*

$$\widetilde{\rho}_{\bar{K}+3,U,\widetilde{U}} = \left(\pi_1 \circ \widetilde{F}_{\omega_{\bar{K}+3}}\right)_* \left(\left(\widetilde{F}_{\omega_{\bar{K}+2}}\right)_* \left(\hat{\ell}_{\bar{K}+1}|_U\right)|_{\widetilde{U}}\right),$$

and let $\widetilde{\rho}'_{\bar{K}+3,U,\widetilde{U}}$ be defined analogously with $\hat{\ell}'_{\bar{K}+1}$ in the place of $\hat{\ell}_{\bar{K}+1}$. Then summing over all possible pairs $(U, \widetilde{U})$ satisfying the conditions above, we obtain

$$\sum_{(U,\widetilde{U})} \|\widetilde{\rho}_{\bar{K}+3,U,\widetilde{U}} - \widetilde{\rho}'_{\bar{K}+3,U,\widetilde{U}}\|_{tv} \leq CL^{-\min\left\{\frac{1}{2}, 1-3\eta, \delta/4\right\}}. \tag{26}$$

Proof. The proof follows closely that of Lemma 17, with $\widetilde{F}_{\omega_{\bar{K}+3}} \circ \widetilde{F}_{\omega_{\bar{K}+2}}$ playing the role of $\widetilde{F}_{\omega_{K+2}}$. For fixed U and $\widetilde{U}$, let $\widetilde{J}$ be as in Lemma 24 for the $\widetilde{U}$ chosen. We consider $\bar{x} \in \widetilde{J}$, and let $(\mathbf{x}, b) \in U$ be such that $\pi_1 \widetilde{F}_{\omega_{\bar{K}+3}} \circ \widetilde{F}_{\omega_{\bar{K}+2}}(\mathbf{x}, b) = \bar{x}$. Then an argument similar to that in Lemma 17 shows that the density of $\widetilde{\rho}_{\bar{K}+3,U,\widetilde{U}}$ at $\bar{x}$ is, up to $(1 \pm CL^{-\min\{\frac{1}{2}, 1-3\eta, \delta/4\}})$, equal to $1/|d(\widetilde{F}_{\omega_{\bar{K}+3}} \circ \widetilde{F}_{\omega_{\bar{K}+2}})_{(\mathbf{x},b)}\partial_1|$. □

A fraction $> 1 - CL^{-\eta}$ of $\hat{\ell}_{\bar{K}+1}$, respectively $\hat{\ell}'_{\bar{K}+1}$, is mapped under $\widetilde{F}_{\omega_{\bar{K}+2}}$ and then $\widetilde{F}_{\omega_{\bar{K}+3}}$ to sets of the form U and $\widetilde{U}$. The $(\widetilde{F}_{\omega_{\bar{K}+3}} \circ \widetilde{F}_{\omega_{\bar{K}+2}})$-images of these parts of $\hat{\ell}_{\bar{K}+1}$ and $\hat{\ell}'_{\bar{K}+1}$ – let us call them $\check{\ell}_{\bar{K}+3}$ and $\check{\ell}'_{\bar{K}+3}$ – are in coupling position. The fractions of $\check{\ell}_{\bar{K}+3}$ and $\check{\ell}'_{\bar{K}+3}$ that cannot be coupled due to unequal weights locally is given by (26). The rest of the estimates are as in the proof of Theorem 2, on choosing $\eta < \min\{\delta/4, 1/3\}$.

References

1. Benedicks, M., Carleson, L.: On iterations of $1 - ax^2$ on $(-1, 1)$. Ann. Math. **122**(1), 1–25 (1985)
2. Benedicks, M., Carleson, L.: The dynamics of the Hénon map. Ann. Math. **133**(1), 73–169 (1991)
3. Benedicks, M., Young, L.-S.: Sinai–Bowen–Ruelle measures for certain Hénon maps. In: Hunt, B.R., Kennedy, J.A., Li, T.-Y., Nusse, H.E. (eds.) The Theory of Chaotic Attractors, pp. 364–399. Springer, Berlin (1993)

4. Blumenthal, A., Xue, J., Young, L.-S.: Lyapunov exponents for random perturbations of some area-preserving maps including the standard map. Ann. Math. **185**(1), 285–310 (2017)
5. Bochi, J.: Genericity of zero Lyapunov exponents. Ergod. Theory Dyn. Syst. **22**(06), 1667–1696 (2002)
6. Bochi, J., Viana, M.: The Lyapunov exponents of generic volume-preserving and symplectic maps. Ann. Math. **161**, 1423–1485 (2005)
7. Chernov, N.: Advanced statistical properties of dispersing billiards. J. Stat. Phys. **122**(6), 1061–1094 (2006)
8. Chirikov, B.V.: A universal instability of many-dimensional oscillator systems. Phys. Rep. **52**(5), 263–379 (1979)
9. Dolgopyat, D.: Limit theorems for partially hyperbolic systems. Trans. Am. Math. Soc. **356**(4), 1637–1689 (2004)
10. Duarte, P.: Plenty of elliptic islands for the standard family of area preserving maps. Ann. l'IHP Anal. Non Linéaire **11**, 359–409 (1994)
11. Gorodetski, A.: On stochastic sea of the standard map. Commun. Math. Phys. **309**(1), 155–192 (2012)
12. Graczyk, J., Swiatek, G.: Generic hyperbolicity in the logistic family. Ann. Math. **146**, 1–52 (1997)
13. Hénon, M.: A two-dimensional mapping with a strange attractor. Commun. Math. Phys. **50**(1), 69–77 (1976)
14. Jakobson, M.V.: Absolutely continuous invariant measures for one-parameter families of one-dimensional maps. Commun. Math. Phys. **81**(1), 39–88 (1981)
15. Kifer, Y.: Ergodic Theory of Random Transformations, Volume 10 of Progress in Probability and Statistics. Springer, Berlin (2012)
16. Lian, Z., Stenlund, M.: Positive Lyapunov exponent by a random perturbation. Dyn. Syst. **27**(2), 239–252 (2012)
17. Liverani, C.: Decay of correlations. Ann. Math. **142**(2), 239–301 (1995)
18. Lyubich, M.: Dynamics of quadratic polynomials, I–II. Acta Math. **178**(2), 185–297 (1997)
19. Lyubich, M.: Almost every real quadratic map is either regular or stochastic. Ann. Math. **156**, 1–78 (2002)
20. Masmoudi, N., Young, L.-S.: Ergodic theory of infinite dimensional systems with applications to dissipative parabolic PDEs. Commun. Math. Phys. **227**(3), 461–481 (2002)
21. Newhouse, S.E.: The abundance of wild hyperbolic sets and non-smooth stable sets for diffeomorphisms. Publications Mathématiques de l'Institut Des Hautes Études Scientifiques **50**(1), 101–151 (1979)
22. Pesin, Y.B., Sinai, Y.G.: Gibbs measures for partially hyperbolic attractors. Ergod. Theory Dyn. Syst. **2**(3–4), 417–438 (1982)
23. Ruelle, D.: A measure associated with Axiom-A attractors. Am. J. Math. **98**, 619–654 (1976)
24. Wang, Q., Young, L.-S.: Strange attractors with one direction of instability. Commun. Math. Phys. **218**(1), 1–97 (2001)
25. Wang, Q., Young, L.-S.: From invariant curves to strange attractors. Commun. Math. Phys. **225**(2), 275–304 (2002)
26. Wang, Q., Young, L.-S.: Toward a theory of rank one attractors. Ann. Math. **167**(2), 349–480 (2008)
27. Young, L.-S.: Decay of correlations for certain quadratic maps. Commun. Math. Phys. **146**(1), 123–138 (1992)
28. Young, L.-S.: Statistical properties of dynamical systems with some hyperbolicity. Ann. Math. **147**(3), 585–650 (1998)
29. Young, L.-S.: Mathematical theory of Lyapunov exponents. J. Phys. A Math. Theor. **46**(25), 254001 (2013)

Communicated by C. Liverani